

Freedom and its obligations

The transformation of Eastern and Central Europe in the past six months presents Europe with a fateful choice between the future and, in one sense, the bad old past. Here is one way in which it can be made.

TUMULTUOUS events in Bucharest, and elsewhere in Romania, in the past two weeks have set a daunting agenda for the Europe of the 1990s. We shall be lucky if the problems now occasioned have been dealt with sensibly before this century (and millennium) are at an end. In at least one respect, indeed, the problems now staring Europe in the face are arithmetically insoluble within one decade or even several. In the countries of Eastern Europe — Poland, Hungary, Bulgaria, Czechoslovakia and now Romania — in which old-style Communist governments were thrown out in the second half of last year, one of the reasons why the people have 'chosen freedom' is that they have seen the previously hidden proof in the West that a centralized state is a self-impoised state. But there is not enough capital about to create Western-style prosperity in Eastern Europe within a mere few years. If Romanians are still digging graves with picks and shovels a decade from now, will their taste of freedom then turn sour?

The further difficulty is that there is as yet no means of telling what will eventually happen in the Soviet Union. Great changes have already taken place, not least the Soviet government's recognition that feeding and clothing its population must take precedence over military expenditure. The election of a pseudo-parliament (the Supreme Soviet) in which dissent is licensed, but in a setting in which the votes lie with the entrenched conservative majority, is an important straw in a wind likely to be further strengthened at the Party Congress due to be held in June. But, for all the daring of the past year, *perestroika* moves at a snail's pace. Price reforms regarded two years ago as essential ingredients of change are now promised only for 1991.

Surprise

The surprise of the past year is the Soviet Union's attitude towards change in Eastern Europe. Mr Mikhail Gorbachev, defining policy as events unfolded, reached a consistent position only with Romania. The countries of Eastern Europe are sovereign states entitled, under the Helsinki agreements of 1978, to self-determination. Even when the intervention of the Red Army in Romania might have been generally welcomed, it did not budge from its bases in neighbouring states. But Gorbachev has also made it clear that he will not countenance the defection of members from the Warsaw Pact nor hankering after

secession from the Soviet Union of its member republics. Lithuania is the cheekiest, but Estonia and Latvia would quickly follow if Lithuania were allowed its way. The Moldavian Republic's wish for *anschluss* with Romania has been similarly well-flagged in the past few months.

Response

How should the world respond to Gorbachev's dilemma? Even those who hold that the yearning for personal liberty is a naturally selected behavioural trait have been given pause by some of the developments of the past six months. Rejoicing at what has happened would have been less sombrely alloyed if the Ceausescu had not been summarily executed on Christmas Day. The ageing dictator's plaintive refusal to recognize the court that sentenced him and his wife to death will haunt the Romanian revolution for some time to come, reminding even Romania's well-wishers of that country's curious deportment in the decades between this century's two European wars. It is also relevant that, in Yugoslavia, the old question of the 1970s, what will happen when Tito dies, has in the event been answered by the return of a political pattern not very different from that of the first decade of this century — balkanization sustained by ethnic rivalry and suspicion.

There are two visions of what lies ahead for Central Europe, one of them a nightmare. The best hope, of course, is that there should be a return to the genteel circumstances of the late nineteenth century, when the Paul Ehrenfests and Ernst Machs of that world found it natural to teach at Prague before moving on elsewhere (to Leyden and Vienna respectively). Sadly, that freedom to move freely was limited, and sustained first by the Austro-Hungarian empire, as bankrupt of a sense of progress as any later tyranny, and just as vulnerable to fissiparous tendencies as is the Soviet Union now. The occasion, not the cause, of the First World War, was Serbian uppishness. The nightmare for Europe now is a return to the state of affairs in which Central Europe's energies are dissipated on fruitless ethnic (really linguistic) quarrels, boundary disputes and economic jealousies. To the extent that this is Gorbachev's worry, it should be shared by the rest of Europe.

How to avoid this chilling prospect? Half-way through 1989, Gorbachev made a case for another conference on European security under the auspices of the Helsinki



Treaty. Events have shown his case to have been strong. Now, luckily, the agenda is clear. Here is what such a conference should aim at:

- A moratorium on boundary changes in Central Europe for some clearly specified period of time, say a decade, and a definition of the circumstances in which they may afterwards take place legitimately. Such a deal would prevent the *anschluss* of Moldavia with Romania and the formal reunification of East and West Germany. But on the Helsinki principle of self-determination, sovereign states may agree on fission (which implies that the Baltic states are Gorbachev's problem, not other people's).

- Reaffirmation that present military alliances (the Warsaw Pact and the North Atlantic Treaty Organization — NATO) will persist for the duration of this moratorium, together with newly agreed deployments of foreign forces on other countries' territory. The plain truth is that Europe is not yet ready for more radical change. But a decade's familiarity with a new regime could transform people's perceptions of peaceful co-existence.

- New guidelines on arms control in Europe. The plan to sign in the second half of June a treaty on conventional arms based on the Vienna negotiations, apparently over-optimistic just a few months ago, is no longer sufficient for the urgent need. Does Europe as now transformed need what will amount to nearly twice the present firepower of NATO, and in the present mix? Is the treaty on missiles of intermediate range (INF) still appropriate when NATO's planned 'modernization' of its nuclear forces seems increasingly a dead duck? Whatever happened to the French proposal for a uniform arms control regime stretching from the "Atlantic to the Urals"?

- An open negotiation of the West's strategic embargo on the export of technical knowhow to Eastern Europe and the Soviet Union, at present decided by fiat by the shadowy Paris-based committee known as CoCom, which has become an expensive anachronism. The original objective, that of denying the Soviet Union and its allies access to US military technology, has long since been subsumed in calculations of the degree to which much new technology, by making civil industry more efficient, may add to economic strength which is the foundation of military preparedness. Now, with the Soviet economy in manifest disrepair, the argument needs standing on its head: will the world be better off if *perestroika* never gets a chance or, alternatively, will high-technology companies in the West rejoice if their products are simply pirated by the now-leaky commerce between West and East? There must be many US software companies that would give their eye teeth for the right to collect \$10 for every illicit copy of their products already in use in Moscow and elsewhere in the Soviet Union.

- An understanding on when elections are free elections. The decade ahead will be riven with endless arguments about the legitimacy of particular European governments, and thus about the circumstances in which they

have been elected. The definition of the franchise is simple enough (but would render some Swiss cantonal governments illegitimate); the problem is to ensure that candidates have been freely chosen and have fought on a level playing field. May it be that the often-surprising willingness of the military powers in the past year that military installations should be inspected in the cause of arms control will pave the way for the international monitoring of elections, perhaps by a European electoral commission?

- The geographical remit of many European institutions needs to be extended. The Council for Europe is one, the European Science Foundation (ESF) another. It is also plain (see page 5) that the new governments of Eastern and Central Europe urgently need assistance in putting their research enterprises back on their feet. They should seize this chance to dispense with the centralized academies of science in the Leninist mould wished on them in the past 40 years, but they would do this the more readily if there were a fund of hard currency, perhaps administered by the ESF, to which researchers could turn for assistance in the short term.

The obvious drawback in this agenda is that it resembles, especially in its first three items, the terms of reference of the Congress of Vienna which, in 1814–15, was intended to put an end to the turmoil of the Napoleonic wars, but which also sanctified the notion that a peaceful Europe requires a delicate balance of military power. Something much like that, of course, has kept Europe ossified for the past 40 years, but the concept has now been shown to be out of fashion. What can replace it?

Model

The obvious model is that of the European Communities, a device by means of which 12 Western European states have agreed to indulge their natural quarrelsomeness only at Brussels, but in the expectation of enhanced prosperity as a result. Eastern Europe needs such a device, but one distinct from Comecon — the economic counterpart of the Warsaw Pact which has the effect of impoverishing its members (to whom Soviet oil is sold too cheaply, but which pay tribute in return). A decade's moratorium on upheaval should allow the Eastern states to fashion something effective along these lines, especially if the European Communities' sensible plan for an investment bank turns out to be effective. Then there could be a deal between the two European economic blocs and the not inconsiderable states that are so far uncommitted.

To ask that Eastern European governments should, while still flushed with recent triumphs, subordinate their inclinations to pan-European causes is to ask a lot, but not unreasonably. The temptation to believe that the new government in Romania marks the end of some process is natural, but mistaken. The truth is that Europe is now poised between two courses of development, one of which points backwards. Will Europe have the patience to choose the other? □

New dangers in East Germany for science

- Change brings materialism, competition
- Initial consensus breaking down

Berlin

THE opening of the Berlin Wall has brought new danger to East German research institutes. Democratic reforms in research institutes are being held up by dissension; meanwhile, pressure created by the opening of the borders is growing every day.

A microcosm of the debate about the way to reorganize research in East Germany can be seen at the Central Institute for Molecular Biology, ZIMB, in German, at Berlin-Buch. Political change has come rapidly to ZIMB. But, surprisingly, exhilaration and joy are not so evident. The subdued reaction might be explained by the peculiar psychology of life in a closed country. "We have lived all our lives within sight of the wall", says Thomas Hanke, a young ZIMB researcher, "It would have driven us crazy to think all the time, 'You cannot jump over it.' So we learned not to think about it."

Hanke, who was born in West Berlin, is one of many who reject the materialism of Western culture. For a lot of East Germans, he says, "real freedom means being able to buy ten Coca-Colas, drink them in public and fire the cans into the corner!" Such views often go along with a desire for reunification and the quick solution this could mean for East German economic problems. But acrimonious debate over whether reunification is the best course has already begun to splinter the popular coalition that rallied behind New Forum to drive the government from power.

The "greatest fear" now is that the coalition might break down and chaos follow, says Werner Ebeling, a professor of physics at the Humboldt University. He and other academics are struggling to find a 'third way' for East Germany that does not involve annexation, confederation, or some other form of reunification with West Germany.

Tom Rapoport, a senior molecular biologist at ZIMB, sees a particular danger for science in the opening of the wall. "It means that we are now competing with the West for good students, postdoctoral fellows and funds. Now we have to be able to offer people something to keep them here, and we have to work on bringing people in from outside."

Rapoport estimates that there are only about 15 groups in the whole of East Germany that come up to international standards in molecular biology and biochemistry, and only a few of these have

publications to prove it. Rapoport has been prodding granting agencies in West Germany and the United States for research support. So far, the response has been "negative or neutral", but he plans to keep "bombarding them" with queries.

If help does not come soon, Rapoport says, "basic research here will die" because East Germany will not or cannot finance it. "We can't wait", he says. Together with colleagues at ZIMB, Rapoport is discussing the establishment of a peer-reviewed granting system in East Germany, assuming that hard currency comes from somewhere. Rapoport recognizes that such a competitive system would shut out at least half the research groups in the country. "Our industry can use them", he says. "It has no good research of its own."

Jens Reich, a molecular biologist at ZIMB who is also a prominent member of New Forum, does not agree. He likens Rapoport's plan to "raw capitalism" with none of the social supports that East Germany must continue to provide if it is to be different from the West. Adopting Rapoport's plan, like adopting the Western economic system, would be a "gesture of capitulation", he says.

Reich would prefer to give junior and senior scientists as well as technical personnel a say in the choice of group leaders and institute directors. He sees a research institute as a parliament, where the leader can continue only when he has the support of a majority of the members. The decision to pursue certain research goals may be taken by a committee, but it must be approved by the institute as a whole.

Rapoport would accept a scientific advisory board with veto power over a director's decisions. But he rejects "letting a cleaning woman decide" how money is spent.

Rapoport admits that he is in deep conflict over the social effects of an achievement-oriented science policy. "We'll dig the grave of a 'third way' by insisting on efficiency", he says. "Personally, I want the social net that Reich advocates. But professionally, I cannot accept it. It is a conflict between feeling and reason."

The debate at ZIMB and elsewhere will have to move forward quickly. Only a few short months remain before the elections planned for May 1990, and reformers still face two daunting tasks — not only must they offer reasonable alternatives to the old system but they must also dismantle

AIDS VIRUS

Dingell reopens Gallo—Pasteur row

Washington

ALLEGATIONS in the *Chicago Tribune* that National Cancer Institute (NCI) scientist Robert Gallo did not tell the full story of the discovery of the AIDS virus in his celebrated disagreement with Luc Montagnier of the Pasteur Institute are to be evaluated by the NCI.

This much is disclosed in a letter from William Raub, acting director of National Institutes of Health (NIH), to John Dingell, chairman of the House of Representatives subcommittee on Oversight and Investigations. At some point, the letter says, Raub and the NIH Office of Scientific Integrity will review the NCI's findings.

The letter is a reply to an earlier letter from Dingell asking whether allegations raised in the *Chicago Tribune* article have been investigated and which continues: "If they have not been investigated, does NIH plan to conduct an investigation . . . ? If so, what office and persons at NIH will be involved in performing the inquiry, and what procedure will be used? When will the inquiry start and what is the estimated time for completion?"

The *Chicago Tribune* article, of the 19 November, was written by Pulitzer-prize-winning journalist John Crewdson (see *Nature* 342, 462; 1989). Among other things the article repeated allegations raised during the Pasteur dispute (eventually settled out of court) that Gallo's identification of the AIDS virus, reported in 1984, rested on the availability of isolates provided by Montagnier.

Dingell's letter to NIH is stern, almost accusatory, in tone, claiming that previous investigations by his staff show that "NIH has turned a blind eye to misconduct by senior scientists . . .". It demands that the new allegations should be "thoroughly investigated . . .".

Raub's reply seems unlikely to satisfy Dingell. It does not say what form the evaluation will take nor whether it has been given the high priority that Dingell would clearly like to see. Nothing is said of when the evaluation will be complete, nor does Raub say whether the NCI evaluation is a response to Dingell's demand for an inquiry or an initiative by NCI. But Raub does say that the *Chicago Tribune* article appears to be erroneous and incomplete in some respects.

Alun Anderson

the old structures of power, which are effectively still in place except at the highest levels. As Reich puts it, "the head has been chopped off but the corpse is still lying in the swamp". It is no coincidence that demonstrations and meetings go on outside working hours. "The courage we saw on the streets has not reached the workplace", he says. **Steven Dickman**

Britain releases green bill

London

A NEW method to measure pollution from Britain's dirtiest factories and legislation on the release of genetically engineered organisms are outlined in the government's wide-ranging environmental protection bill, published just before Christmas. The bill, popularly known as the 'green bill', deals with a collection of issues that have accumulated at the Department of the Environment over the past decade.

The bill introduces integrated pollution control (IPC) which will require the effect of pollutant emissions to air, land and water to be assessed together, rather than separately as at present. Some 3,500 industrial sites will be covered by IPC and will have to obtain and pay for authori-

COMPUTER MISUSE

Strict law proposed

London

TOUGH new penalties for computer crime are proposed in a private member's bill presented to the British parliament. The bill, introduced by Conservative Member of Parliament Michael Colvin, would define three new crimes. The first is a 'basic' offence of unauthorized access to a computer system which would be punishable by a maximum fine of £2,000 and/or six months' imprisonment. Two more serious offences, unauthorized entry into a system with intent to commit a crime and unauthorized alteration of computer-held information (including the circulation of disks carrying computer 'viruses' and other destructive programs), would carry a maximum penalty of five years' imprisonment.

Colvin's bill follows the recommendations of an earlier Law Commission report but doubles the penalties for the basic offence of unauthorized entry, or hacking. The Law Commission report arrived too late for its suggestions to be presented as a government bill this year, and Colvin believes that the problem of computer misuse demands urgent action, if Britain is not to become an "international centre for computer crime". Estimates of the financial cost of computer crime in Britain vary between £400 million and £1,000 million a year. Colvin says the decision to propose the bill was influenced by the recent outcry over the distribution of a 'trojan horse' program, disguised as information on AIDS, to thousands of British computer users.

Private members' bills have great difficulty in becoming law because it is hard to find time for them in the parliamentary schedule. But Colvin is optimistic. The bill will be debated in detail at its second reading in the House of Commons on 9 February.

Peter Aldhous

zations to operate from Her Majesty's Inspectorate of Pollution (HMIP). Authorizations will demand that factories use the "best available techniques not entailing excessive cost" to render emissions harmless. Contravention of the specific conditions contained in authorizations may be punished by an unlimited fine. A second tier of less-polluting factories would not be subject to IPC, but the government intends to give local authorities new powers to control industrial air pollution.

Chris Patten, Secretary of State for the Environment, claims that the bill creates "foundations for pollution control well into the next century". But there is concern that HMIP will not have the manpower to police the new IPC controls. A recruitment drive is under way to fill 33 vacancies and it is hoped that recent pay rises will help to relieve a crisis of morale within the inspectorate. At a press conference, Patten said that the current recruitment drive may precede a strengthening of HMIP.

New controls on the release of genetically manipulated organisms are also contained in the bill. They will implement some of the recommendations from the Royal Commission on Environmental Pollution report, published in July (see *Nature* 340, 84; 13 July 1989). Those seeking to release genetically manipulated organisms will need to provide an assessment of the environmental hazards posed by the release, and must then wait for consent to be granted by the Health and Safety Commission (HSC) and the Secretary of State for the Environment (in some cases, the Ministry of Agriculture may also be involved).

But, as indicated in a government consultation paper released in June, some releases may be exempt from the system of consents. The precise criteria by which exemptions are to be granted will become clear only as the bill is debated in parliament. Another area to be clarified is the extent to which the content of applications for consents will be made public.

The new regulations will complement existing legislation that stresses human health and safety, rather than environmental protection. That legislation is enforced by the Health and Safety Executive (the operating arm of the HSC) with help from the Advisory Committee on Genetic Manipulation. Under the new proposals, the advisory committee would continue its function, and a new committee would advise the Health and Safety Commission and the environment secretary on the environmental implications of the release of genetically manipulated organisms. A single consent for each release would then be granted on the basis

Scientist as science minister

New Delhi

PRIME Minister Vishwanath Pratap Singh, whose National Front party came to power in the recent Indian election, pulled a surprise at the end of last year by appointing Professor M.G.K. Menon as the minister for science. Menon had held scientific posts in Congress-led governments since 1971 and was science adviser to the former Prime Minister Rajiv Gandhi.

Menon was an associate of Homi Bhabha, father of India's atomic energy programme, and is a well-known cosmic-ray physicist. He said the appointment came as a surprise. He is the first scientist to become science minister in independent India. His predecessors were all diplomats or elected politicians. In choosing Menon for the job, Singh, himself a physicist, was guided by his philosophy that science is non-political and that a professional is needed to steer India's vast scientific establishment.

As a member of the planning commission since 1983, Menon was in charge of formulating five-year science plans and allocation of funds for various departments. Before that, he had served as secretary in the departments of electronics, defence research, science and technology and environment. He was chairman of the UN secretary general's committee for science and technology for development. He has been president of the International Council of Scientific Unions since 1986.

Menon said his first task would be to integrate the multiplicity of departments and improve links between research, industry and education.

He expects some changes in research priorities but has ruled out any major shift. "We cannot pull back programmes already committed", he said, "but we can always review the phasing of these." K. S. Jayaraman

of both human health and environmental considerations.

Also included in the 'green bill' are new rules on the dumping of waste at sea and waste disposal by local authorities and increased penalties for littering. While welcoming some of its proposals, the Labour opposition and environmental pressure groups have attacked the bill for being a "hotchpotch" of unconnected measures, and for its proposal to split the government's advisory body on conservation issues, the Nature Conservancy Council, into separate agencies for England, Scotland and Wales. Also criticized is the failure to include controls on the emission of carbon dioxide. Future government policy on this and other environmental issues will become apparent in a white paper (policy document) to be published next autumn. **Peter Aldhous**

Survey of science — a new decade begins

At the beginning of the 1990s what more astonishing sight than that of Eastern Europe, a year ago still isolated behind the iron curtain, standing on the verge of reintegration into the common European home of science. The changes in Europe and the Soviet Union, as well as the rapidly-growing importance of international cooperation (and competition) in science will set many of the big science policy issues of 1990.

The changes in the East

Europe

DESPITE the political reforms, material relief for Eastern Europe's impoverished scientific community has yet to arrive. In **Poland**, for example, where the 'semi-democratic' elections last June flooded the new two-chamber parliament with scientists and scholars (many of them former dissidents), Jan Janowski, the deputy premier with responsibility for science and technology, announced recently that the "coffers are empty" as far as applied research is concerned.

Research institutes that do not come under either the university network or the Academy of Sciences, he said, will sink if they cannot negotiate contracts with industry. Dozens of institutes could close in the next few months and research scientists seem likely to become the first sector of the community to seek the planned new unemployment benefits *en masse*.

Throughout Eastern Europe, scientists and university lecturers are generally paid at a far lower rate than trained industrial workers. Any wage rises in the coming year are unlikely to do more than mitigate the effects of soaring inflation. Foreign debts mean a lack of hard currency for foreign journals and equipment, and there are fears that foreign colleagues who have been generous with gifts and subscriptions in the past may naively believe that with the coming of democracy the problems of their Polish, Czech, Hungarian or East German colleagues will miraculously be solved. In fact, no immediate improvement can be expected, and, with relaxation of travel restrictions a major brain drain to the West is likely.

Hungary is already losing trained scientists at the rate of ten a day while East Germany is dangerously short of medical specialists. Not surprisingly, a number of eminent visitors to the West during the past few months have stressed the need not for charity but for an infusion of Western knowhow that will help create jobs for scientists and engineers.

Western participation in joint enterprises is now seen both as an important step to economic recovery and as a means of bringing Eastern Europe back into the world scientific and economic community.

East European participation in inter-

national scientific projects will almost certainly increase (East Germany, in particular, has its sights set on CERN, ESA, and Eureka). *Glasnost* will mean lifting secrecy on environmental issues, and new cross-border cooperation to deal with specific problems. In **Czechoslovakia**, Civic Forum has declared its opposition to the controversial Danube hydro-electric scheme, and has approached the Hungarians about creating an Austrian-Czechoslovak-Hungarian wetlands reserve.

In the formerly hard-line countries of **Bulgaria**, **Czechoslovakia** and **East Germany**, rehabilitation of independent-minded scientists and scholars expelled from their positions or censured by the Party is already under way. East Germany is revising the statutes of its Academy of Sciences, while **Poland** will almost certainly end the practice by which the learned secretary of its academy is a government appointee. And Czechoslovakia has set up a commission to enquire into allegations of abuse of psychiatry for the repression of political dissidents.

A possible source of funding for science and education throughout Eastern Europe would be the resources set free by defence cuts. 'Conversion' of military industry to the civilian sector is a current buzz-word in the Soviet Union, but it is unclear just how much this will affect the smaller Warsaw Pact countries. Unlike the Soviet Union, where military and civil research have always been sharply distinguished, the East European countries have encouraged the transfer of technology from the military to the civilian sector. And, although East Germany has the second-largest army in the Warsaw Pact, East German diplomats claim the country has virtually no military industry.

Romania, which had its breakthrough to democracy at the very end of the year, faces a particularly urgent problem. Its late 'Conducator' Nicolae Ceausescu and his wife the "world-ranking scientist and academician, Comrade Dr Elena Ceausescu DSc", during the past two decades had virtually annihilated fundamental research in Romania and yoked all applied research to the fulfilment of politically inspired planning. A massive rescue operation by Western science would seem to be in order. □

Soviet Union

TURMOIL, even constructive turmoil, is not good for science, and there is every

likelihood that the excitements of 1989 will be projected into the year ahead. To be in or near Moscow is absorbing, but inimical to peace of mind.

Lobbying has already begun for the elections to the Soviet Academy, likely in the spring, with Yu. A. Osipyan, the solid-state physicist, mentioned frequently as a likely successor to Gurii Marchuk, the present president. Whether election will bring reform of the academy's cumbersome organization is another matter.

In the year ahead, and perhaps for the first time, there will be a restraint on Soviet research that was hardly ever apparent in the past — a lack of funds. True, hard currency has always been in short supply, whence the appalling general shortage of up-to-date equipment, but now there are signs that even Aeroflot tickets to the West are hard to come by even when there is a willing host waiting at the destination. The chances are that the salaries of researchers will fall even further behind those of, say, taxi-drivers. □

International collaboration

US politicians no longer automatically say that American scientists must do big research projects on their own; whether it be the Superconducting Super Collider, the human genome or the Space Station, the cry is now for international collaboration.

But the problem is always that politicians are attracted to international collaboration because it is a way of saving US money. The moment it is suggested that foreign partners might themselves benefit from collaboration, then suddenly the talk is of others gaining a 'free ride' on US technological expertise.

Much of this is misunderstanding. US politicians do not understand that in most collaborative projects, foreigners participate by invitation because the US partners want access to foreign technical expertise. Times have changed — international collaboration is not a subterfuge by which foreigners gain the benefits of US science and technology at low prices.

The Space Station is the largest of all the collaborative projects. But 1990 may be the year in which the international partners finally abandon attempts to collaborate with the United States in space. When US

proposals to scale back the space station, achieving essentially the same capabilities but a few years late, filtered down to the European Space Agency and Japan's Science and Technology Agency — and to the US Congress — all were furious.

Unless NASA can regain the confidence of its partners and of Congress next year, the future of the space station — and the missions to the Moon and Mars which could depend on the space station as a base for developing a lunar outpost — looks bleak.

Human genome project

An international effort to sequence the human genome is just taking off. Led by James Watson, a man hugely popular with Congress, the genome project's budget was almost doubled this year to about \$60 million but its goal of a \$200-million annual budget within two years seems unreachable.

Watson has already caused consternation in **Japan** with his suggestion that countries failing to contribute adequate finance to the human genome project should be denied access to the fruits of the research. Although such a suggestion is clearly unworkable for a multi-national project, Watson plainly has a legitimate complaint. While the **United States** and the **United Kingdom** have come forward with donations for the Human Genome Organization (HUGO), Japan has failed to scrape together its \$300,000 contribution. That has only served to reinforce the widespread perception that Japan is not serious in its professed desire to contribute more to basic research. But several Japanese ministries now see a chance to make their names in the project.

The Superconducting Super Collider is a different matter. Japan (and several other countries) would prefer not to be asked to help on the grounds that a refusal often offends. But despite recent advice, Japan will not learn how to say "No", but will follow along late and unwillingly.

Europe

The United States and Japan are not at the heart of all collaborative research efforts. Leaving Japan completely out is a European collaboration in microelectronics known as JESSI. Five European countries and three multinational companies decided to back the project with an eye on the Japanese, who they fear will dominate the strategic market in semiconductor chips by the end of the next decade. But like the space programme, JESSI — which stands for Joint European Submicron Silicon — is drawing criticism for pouring public money into an area best left to private enterprise. The United States may eventually become part of the team. The big decision is whether JESSI will link up with a similar US effort called Sematech. □

Cold fusion

The scientific event to be proud of in 1989 was not the announcement by two Utah researchers — Stanley Pons and Martin Fleischmann — that low-temperature fusion-in-a-bottle was possible but the

AP

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

The non-event of 1989?

world-wide activity that it stimulated. The international scientific community showed it could take on board totally novel scientific ideas, digest them and despatch them, all with considerable speed and relatively good humour.

Cold fusion has come and gone in eight

months (23 March was the date of the press conference at which first word of cold fusion reached the world, November the date of the Department of Energy report which virtually ended interest in the phenomenon).

Most of cold fusion's critics are already back at whatever they were doing earlier. Pons and Fleischmann apparently remain undeterred and a few supporters remain. A group at Texas A&M are now the most vocal, but by their own admission they find only 'anomalous heat output', and absolutely no evidence for any nuclear products — no helium, no gamma rays, no neutrons — and the tritium that has been found in some of their experiments is definitely not being produced in a high-energy process.

Occasional claims of success are heard from elsewhere in the world — as recently as December in Japan. But most Japanese scientists are sceptical about the effect. Research continues in India at the Bhabha Atomic Research Centre in Bombay. The director, P.K. Iyengar, believes cold fusion to be the source of the enormous amount of tritium produced in his experiments in which deuterium gas is pumped into titanium under pressure. He claims that US scientists are convinced that cold fusion can take place but are keeping their results secret. □

Space

Voyager's last hoorah at Neptune last year left the **United States** devoid of any active spacecraft with the exception of Magellan, sent on its way to map by radar the Venusian surface next August. But a resurgence of space-based astronomy and planetary exploration is on the way. Galileo was sent off to Jupiter in October. COBE, the long-awaited Cosmic Background Explorer, was launched in November. This year the Hubble Space Telescope is still due for launch in March. The next couple of years look good too, with a strong programme in astronomy. And a full fleet of launch vehicles is now in business: the space shuttle, along with Titan, Atlas and Delta rockets.

Beside the **Soviet Union** (whose own shuttle has flown only once — unmanned — and seems unlikely to fly again) only **France** and **Japan** have hopes of grand achievements in space. France has the more immediate hopes with manned space flight its first objective. France has always been the strongest supporter of the European Space Agency (ESA)'s Columbus contribution to the Space Station and is going to provide 45 per cent of the cost of the Ariane 5 rocket needed to launch the Hermes shuttle.

This year, France will spend FF1,812 million on the Ariane rocket launcher and FF737 million on the European shuttle Hermes and contributions to Columbus.

But Hermes' costs are spiralling and partners elsewhere are growing restive. The election of a Frenchman to the post of director-general of ESA would undoubtedly help — Reimar Lüst leaves in October. But with the United States increasingly uncertain about its commitment to the space station, the need for Hermes is not clear. And if Hermes is unnecessary, then Ariane 5 is not needed either.

West Germany has so far managed to reduce support for Hermes, Ariane and Columbus without endangering the projects. But this year Research Minister Heinz Riesenhuber will have to decide how to raise the next big contribution to the ESA projects, due in 1991. If he takes it out of existing programmes, as he has promised not to do, there will be a tremendous uproar. If he makes further cuts in the ESA budget, he will make West Germany look like a miser in France's eyes.

Japan's plans for space greatness are more distant — but manned space flight is one of its objectives, according to the grandiose official plan released last year. But for the time being it is David, rather than Goliath, that is calling the shots. A neat slingshot manoeuvre should next year enable Japan's low-budget Institute of Space and Astronautical Science (ISAS) to send a satellite to the Moon.

But if Japan is to get serious about space exploration then someone is going to have to bite the bullet and give ISAS access to really big payloads by forcing it and

AP

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Just business as usual

its giant rival, the National Space and Development Agency (NASDA), to work together.

India had difficulties with its space efforts last year. Two test rockets were lost in the sea in succession. One consolation was the successful test of the 130-tonne booster stage of the Polar Satellite Launch Vehicle to be launched in 1991. □

Projects and budgets

In the **United States**, several big projects passed milestones last year. The Stanford Linear Collider (SLC) finally produced some physics results, culminating in the announcement that the number of particle 'families' had to be three. And the Superconducting Super Collider finally began to burrow into the ground. But it has been such a long and painful gestation that the birth was not attended with wild celebrations: many still think the project is a big sink of scarce research money. And there are still problems with the super-conducting magnets.

Defence expenditure is certain to be cut this year as East-West relations improve. That could hurt DARPA, the Defense Advanced Research Projects Agency. DARPA does more than it is given credit for of the kind of work that elsewhere would be supported by a Department of Industry or MITI. Among them are its projects on computer networks, parallel computing, memory chips and neural network computing. If DARPA is weakened, such support is unlikely to come from elsewhere.

A major disappointment in the United States has been the new administration's

Environment

1989 was the year that politicians everywhere turned green, tripping over one another in their haste to organize international conferences on global change. British Prime Minister Margaret Thatcher, who a year ago would have seemed a very unlikely environmentalist, came in first with the London conference on the ozone layer in March. François Mitterrand had his chance next, when he hosted the summit of industrialized nations in July. And Japan followed with its conference in September. Definite measures did not emerge from any of the conferences, beyond support for the Montreal agreement to reduce the production of chlorofluorocarbons (CFCs).

This year it is the turn of the United States with a conference on global warming announced by President George Bush at the recent Malta summit. The United States usually makes its big announcements at conferences it has organized itself — but what can it possibly come up with?

Little more can be done regarding CFCs. Industry everywhere is agreeing to phase out CFC use even faster than the Montreal agreement demands. Worrying exceptions do remain in the developing world, principally China and India.

The big question for this year is whether the United States will call for (or even consider) international agreements to reduce

emissions of carbon dioxide and thus slow global warming. For most nations, cutting carbon dioxide emissions means raising the efficiency with which fossil fuels are consumed, principally in automobiles and power stations. (An exception is Brazil where burning of Amazon forest is also an important source of carbon dioxide.)

But the barriers to improving energy efficiency are high. In the United States, which wastes more energy than any other nation, even the most up-beat environmental lobbyists say that there is only a 50 per cent chance that legislation to improve drastically the fuel efficiency of automobiles will pass through Congress this year. The automobile industry lobby is simply too strong.

In the United Kingdom, action at power stations seems barred by the decision to privatize the electricity industry, planned to begin in March 1990. One of the casualties of that decision is the resignation of Lord Marshall from the Central Electricity Generating Board in protest at the decision that nuclear plants should not be privatized.

Japan has a quite different reason for opposing international agreements to reduce carbon-dioxide emission levels — it already produces far less carbon dioxide per head than other industrial nations. Japan can rightly argue that nations elsewhere should simply be imitating the levels of energy efficiency it has already attained. □

lack of urgency in appointing the top personnel of the science agencies. A new director for the National Institutes of Health has yet to be announced, even though the previous director left six months ago. The new presidential science adviser, D. Allan Bromley, has so far failed to have a visible impact in the White House, raising fears that after neglect in the Reagan era, the job cannot recover the importance it had in earlier administrations. This year will show whether Bromley can live up to his reputation for leadership.

This year, **Japan** may finally be at a turning point in it is support for basic research. Although Japan's competitors may not know it, the government has been practising austerity for a decade, trying to control its deficit. This year, the Ministry of Finance should at last be able to permit more flexibility in budget allocations. If expansion comes, then the Human Genome Project is likely to find substantial support at last.

The psychological impact of Japan's first tiny step for mankind — the opening of the international Human Frontiers Science Program in Strasbourg — will become apparent this year. The budget for the programme is very small, just \$20 million. What is important is that it is the first time Japan has supported research that does not necessarily involve Japanese

researchers. If it is successful, it could help Japan's research system to become more international, more open and more willing to reward talent.

Last year was another good year for **West German** research budgets. The Research Ministry overcame the threat of a budget freeze early in the year and even managed to salvage a small increase. But despite growing budgets and general material well-being in West German research institutions, researchers often worry over how half-heartedly their work is supported by politicians. As this year is an election year, real intentions will be tested.

A change of government in 1990 would certainly affect science. The opposition Social Democrats are likely to call for more environment research and less space and nuclear research.

Discussion of export restrictions took centre stage in West Germany earlier last year, with threats of US sanctions in the aftermath of the Libyan poison gas scandal. But now the issue has been turned on its head by the developments in East Germany. One of the first *de facto* signs of reunification will be the transfer of high technology in the form of computers and industrial equipment to East Germany.

For the first time, Chancellor Helmut Kohl has a strong hand in dealing with the

United States on this issue. In the same way that Kohl declared a ten-point plan for reunification without consulting the allies, he may now openly flout the CoCom rules as a way to help East Germany's ailing industry. By the end of the year, CoCom seemed to have recognized the potential for conflict during 1990 by proposing that restrictions on certain exports to Eastern Europe — but not the Soviet Union — should be lifted. Whether that will be workable is anybody's guess.

In **France** researchers had an average year. CNRS and INSERM (the medical and health research institutes) received modest increases in their budgets — mostly for staff and better working conditions at the bench. University research has been given a boost, and a plan to double the number of university entrants within ten years was declared. But the jerry-built universities are already falling down and cannot handle present capacity — up by 40,000 from last year.

Construction started on the \$434-million European Synchrotron Radiation Facility (ESRF), 14 years after it was conceived within the European Science Foundation. At the European particle physics laboratory (CERN) on the Franco-Swiss border, the Large Electron-Positron Collider (LEP) started running and is expected to produce one million Z^0 particles a year.

In the **United Kingdom**, this year will be one of upheaval for the universities. For the first time, universities must bid against one another for teaching grants from the Universities Funding Council for the academic year 1990–91, stating how many undergraduates they wish to take, and giving an 'offer price' per student in each subject. Universities remain in the dark about the new bidding rules, and the consequences they will bring for the relative strength of various institutions.

Support for research is on the rise after a decade of impoverishment. In November, £61 million was added to the science budget for next year, an increase of 27 per cent over the previous year. although the five research councils do not yet know how the secretary of state, John MacGregor will distribute the money. The five research councils may in any case soon be one as MacGregor is considering a plan to create a single National Research Council.

In **India** the new government is biased toward small science and does not favour large-scale research for the sake of prestige. The watch-words are self-reliance, import substitution and cuts in technology import — bad news for multinationals that so far have had a free run.

Launching a human genome sequencing project, collaboration with the United States on the Superconducting Super Collider, and signing the Paris Convention on intellectual property rights are some of the major issues on which the

new government has to take decisions. There are plans to grant greater autonomy to government research institutions and to change the official secrets act so that research in government laboratories (80 per cent of the total) will become more transparent and accountable to the public.

In **Australia**, people will be as much concerned in 1990 as last year with the ramifications of dawkinsization — the radical reorganization of higher education and research being pushed through by the Minister for Employment, Education and Training, J.S. Dawkins. The objectives of

the new policy are wider access to higher education and concentration of research funds in centres of excellence. The overall budget is growing at a healthy rate.

But the methods have caused much distress, and will continue to do so. Some academics resent (and resist) the enforced merging of institutions of different character. Others are deeply suspicious of the interference from the centre that has followed to drive for efficiency and accountability. Many academics believe that it will "all be different after the election", but they could be wrong. □

Ethics and research

Sex, hearts and brains

SCIENCE may be universal but the ethical issues that science generates are not. Every culture has its own worries — the poor old Anglo-Saxons are still unable to talk openly about sex it seems. The governments of the **United States**, the **United Kingdom** and **West Germany** last year refused to carry out surveys of sexual behaviour that would greatly help to predict the spread of the AIDS virus. "Too intrusive", said the British. In the end the Wellcome Foundation picked up the £900,000 bill for the UK survey. The French reacted differently — they approved a national survey of sexual behaviour without embarrassment.

In **Japan**, for a mixture of cultural and religious reasons, the general public has difficulty in accepting brain death (rather than the cessation of the heartbeat) as the death of a person. The result is that no heart transplant operations are performed. 1990 could see this change, with a public consensus for the acceptance of brain death gradually emerging.

The practice of abortion, in contrast, presents Japanese with no specific moral dilemma. The same is true in **India**, where abortion is legal and encouraged as a population control measure. India will this year establish a human fetal tissue bank under a \$400,000 grant from the Department of Science and Technology. Operating from the All India Institute of Medical Sciences in New Delhi, the bank will supply fetal tissues to researchers in developmental biology and organ transplantation, particularly those involved in the transplantation of neural tissue.

The **United States** has gone to the opposite extreme. The smooth appointment of senior administration officials was disrupted last year by an apparent need for candidates to demonstrate anti-abortion views acceptable to the administration. And a ban was renewed on research involving the transplantation of tissue from aborted fetuses into human recipients. But towards the end of the year, the pro-abortion lobby (or 'pro-choice' lobby as it likes to be known) became more vocal.

Candidates with strong anti-abortion views fared badly in state elections.

Vigorous opposition to the renewal of the ban on fetal tissue transplant research can be expected this year as groups representing people who might ultimately benefit from such research — diabetics and people with Parkinson's and Alzheimer's diseases and with AIDS — move into action.

The **United Kingdom** has yet to give serious consideration to the issue of research on fetal tissue transplantation. But it took the middle road with an exhaustive report on the related ethical question of research on embryos intended to improve *in vitro* fertilization techniques. (Such research is permitted in the United States although the anti-abortion lobby would like to see it banned.)

Members of Parliament will this year be offered a choice, voting on a bill which contains two alternative clauses, one banning all embryo research, the other allowing closely regulated research on human embryos up to 14 days after fertilization. Members will be allowed to vote as their conscience dictates and will not be expected to take a party line.

Animal rights

Animal rights activists continued to strike at research laboratories in the United States, burning down buildings at the University of Arizona. **France** saw the first such action for many years when nearly 100 animals, including 38 monkeys, were 'liberated' from two INSERM laboratories in Lyons. The monkeys were eventually returned, but animal vivisection had become an issue overnight in formerly liberal France.

But when the **United States** attempted to prepare tough new regulations for the care and use of animals in laboratory experiments, the research lobby struck back. In the end, the federal government gave in to pressure from researchers and pharmaceutical companies who argued that the cost of implementing the regulations, would inhibit research.

A good compromise between researchers and those concerned with animal rights seems to have been worked out in Cambridge, Massachusetts, where a local ordinance to govern the care of laboratory animals was enacted.

Genetic engineering

In the **United States** controversy switched last year from the release of genetically engineered organisms to gene therapy. After surviving an arduous series of examinations and re-examinations by regulatory bodies, W. French Anderson and his colleagues at the National Institutes of Health finally received permission



Gene therapy survives red tape

to insert a bacterial gene into humans as a marker in an anti-tumour therapy.

Elsewhere in the world, governments made efforts to provide new regulations for biotechnology. In the **United Kingdom**, it became a criminal offence to release genetically manipulated organisms without first notifying the government's Health and Safety Executive.

West Germany continued to find the regulation of biotechnology way beyond its competence. Last year, it once again failed to set up clear legal guidelines for genetic engineering, despite pressure from the pharmaceutical industry and university researchers. The government did make progress in sketching the outlines of the new 'basic law' for genetic engineering but it left the difficult details.

Researchers in Cologne received permission to carry out the first release of genetically engineered organisms in **West Germany**. But their plan to grow 40,000 pink petunias carrying recombinant DNA was postponed until this year. Environmentalists oppose the release because it sets a precedent.

Europe put the brakes on biotechnology too last year. The European Commission announced in August a 15-month moratorium on the use of bovine growth hormone (BGH), a peptide hormone produced using genetic engineering. BGH increases milk production in cows with minimal side-effects, say the manufacturers. The Commission said it needs more time to evaluate scientifically possible risks before approving the sale of BGH. □

Fraud and misconduct

This year surely cannot be like last with its string of US congressional subcommittee hearings on misconduct and conflict of interest in science, none of which appeared to produce useful results in proportion to the enormous time and effort they consumed. The main conclusion of the pair of hearings concerning the case of David Baltimore was surely that there was no conclusion — except perhaps that a congressional hearing is not a suitable place for politicians to try to debate immunology with a Nobel laureate.

But this year may hold some surprises. Even the Baltimore case may not yet be dead. Investigations are continuing of laboratory notebooks containing data upon which rest some of the conclusions of the disputed *Cell* paper. But the science involved is so esoteric that the subcommittee may find a forceful conclusion can never be drawn.

A clearer and simpler case of alleged fraud came from India last year. A researcher from Punjab University allegedly polluted literature on the geology of the Himalayas for over 20 years by passing off museum and other samples as fossils he had recovered on field trips.

Nuclear energy, waste and weapons

1989 was the year in which the **United States** began to face up to the colossal problems of the accumulation of waste from decades of nuclear weapon production. Under the new leadership of Admiral James D. Watkins at the Department of Energy, the gargantuan cost of cleaning up nuclear contamination has finally been counted — 30 years and \$200,000 million is required, \$19,000 million is needed over the next five years just to get started.

The department's plans to build a repository for high-level nuclear waste from commercial nuclear power plants at Yucca Mountain, Nevada, is also in difficulty. Completion of the project is now impossible before 2010, seven years past the original deadline. And if further study finds the Yucca Mountain site unsuitable for the waste, the problem will be back in the hands of Congress.

In the **United States** there have been no new orders for nuclear power plants since the Three Mile Island accident in 1979, and all orders dating back to 1974 have been cancelled. Last year, voters in Sacramento, California, became the first to shut down an operating nuclear power plant.

Another year passed for the still-unlicensed nuclear plants at Shoreham and Seabrook. Watkins has argued forcefully for both facilities but New York has decided to purchase the Shoreham plant from its owners for one dollar and then to assume the costs of shutting it down. Watkins has called this "one of the most foolish deals in the nation's history". Seabrook could be the one glimmer of good news for the nuclear industry, although so far it is a small glimmer. After the utility that owns Seabrook declared bankruptcy, the Nuclear Regulatory Commission finally agreed to find a way to license the plant.

The British nuclear power industry faces a year of reorganization. In January, most existing Magnox and advanced gas-cooled reactors will come under the control of two new divisions of the Central Electricity Generating Board (CEGB): which will eventually be converted to separate government-owned companies. Further development of **Britain's** nuclear programme is to be shelved until 1994. Plans to build three new pressurized-water reactors are on hold and the development of new reactor designs will be delayed.

West Germany's plans for nuclear power generation are becoming ever less grand. The controversial nuclear fuel reprocessing plant at Wackersdorf has been abandoned. And with the completed but unlicensed fast-breeder reactor at Kalkar unlikely ever to go on line, nothing remains of the plan to complete West Germany's nuclear fuel cycle.

Now that nuclear energy opponents have won the battle against Wackersdorf, they are likely to begin efforts to ban nuclear energy in Europe. The strong showing of Green groups in European Parliament elections in June 1989, bodes ill for nuclear programmes all over Europe.

Even in **Japan**, long the most stalwart supporter of the virtues of high technology in all its forms, the anti-nuclear power movement is growing ever stronger. The Japanese government and power industry have tried to reverse the trend through publicity campaigns but to little avail. A pump failure at one of Japan's nuclear power reactors, the Fukushima No 2 plant caused considerable public alarm. The reactor has since been out of action.

The issue is likely to come to a head this year as the government and industry hope to open part of a huge complex in Aomori Prefecture in 1991 for storing radioactive waste and for reprocessing and enriching nuclear fuel. □

From Peter Aldhous (United Kingdom), Alun Anderson (United States), Peter Coles (France), Steven Dickman (Germany), KS Jayaraman (India), David Lindley (United States), John Maddox (Soviet Union and Australia), Christine McGourty (United States), Vera Rich (East Europe), Seth Shulman (United States) and David Swinbanks (Japan).

Research on human embryos

SIR—I, for one, welcome the recent decision by Dr Louis Sullivan, Secretary of Health and Human Services, to extend the ban on the use of US federal funds to support research on human embryos and fetuses obtained by elective abortion. The statement in your leading article (*Nature* **342**, 104; 1989) that “the question of whether abortion should continue to be allowed . . . is a separate question” would obviously be incorrect if elective abortion were illegal. Using tissue obtained from an illegal procedure would presumably itself be illegal. The question of whether it is acceptable to use the abortus for research or other ends arises only because abortion is legal. In fact, if legality were the only consideration, ‘anti-abortionists’, as you label them, would have to concede that the use of fetal tissue derived from a legal procedure should itself be legal.

Some of us, however, believe that the legality of abortion is irrelevant to its morality, in which case the resolution of the fetal tissue question does not derive strictly from legal issues but must involve determination of the morality of elective abortion. If, as some of us believe, elective abortion can only rarely be defended on moral grounds, then further profit from such acts, either by individuals or by society, is also immoral. The same objection can be, and has been, applied to the use of data from Nazi experiments. However, unlike the Nazi medical data, which are historical, the use of aborted tissue is feasible only if there is a continued practice of abortion.

Thus, my objection to the use of fetal tissue stems from the perceived immorality of the act that leads to its procurement. Is it not ironic that the two properties of human fetal tissue that make it so desirable for medical research and therapy, namely, that it is undeniably human and that it has tremendous developmental potential, are two of the same properties that bring pro-life advocates to the defence of these lives?

On the other hand, if elective abortion can be justified on moral grounds, then let us press boldly onwards with complete freedom for abortion, whether for gender selection or to provide fetal tissue for research or treatment. In fact, if abortion can be justified on moral grounds, I would argue that it would be immoral not to use fetal tissue for all possible benefit, including direct consumption for nutritional purposes. I am somewhat surprised that no movement has appeared among pro-abortionists in support of abortuses as a food source. Why waste this valuable resource? Some might object that this borders on cannibalism, but that would be the case only if the victims were human. Besides, it is not obvious what moral or

legal difference, if any, exists between the maintenance of health and life through human tissue transplantation as opposed to its direct consumption.

KEITH A. CRUTCHER

*University of Cincinnati Medical Center,
College of Medicine,
Department of Neurosurgery,
231 Bethesda Avenue,
Cincinnati, Ohio 45267-0515, USA*

Women in science

SIR—As a woman scientist, I should like to comment on the situation of women in West German science.

“If a woman has a special gift — which does not happen very often — I do not think it right to refuse her the chance and means to study . . . but such a case must always be regarded as an exception” (Max Planck, 1897¹). In his time, Planck was considered a progressive man to tolerate women in the scientific field. Only in 1908 were women allowed to study in German universities. In 1923, the first woman was nominated professor.

In 1989, women are still rare in the natural sciences. At universities as well as at research institutes (such as those of the Max-Planck-Society) women are mostly represented as students, technical assistants and secretaries. There are almost no women as team heads or independent researchers. In 1987, only 6.8 per cent of all researchers in mathematics and natural sciences working in institutes of the Max-Planck-Society were women². Higher positions are without exception occupied by men. And among the 133 active participants at the spring meeting of the German Physiological Society this year, only three scientists were women.

There are several reasons for the poor representation of women in German sciences. One is the lack of tradition (with some famous exceptions³). Another, more important, reason is the still patriarchal and authoritarian system of German universities and other scientific institutions. Then there are economic problems. The German tax system makes it very unattractive for both partners in a marriage to work. Part-time jobs are almost impossible to get in science. The belief is that if one does not spend 12 hours a day in the laboratory one is not seriously dedicated to science. In this context, the notion that women could simultaneously be traditional mothers, traditional wives and productive scientists seems entirely absurd.

Women are allowed to get a PhD, but who assures them they will find a job afterwards — particularly if they intend to marry and have children? They are confronted with the alternatives of family or career. In fact, most of them leave their

jobs when they become mothers.

Paradoxically, Latin countries with ‘macho’ traditions offer more opportunities for tenure for professional women in almost all academic fields. It is a pity that in a technologically developed country such as Germany, people waste their talent and energy because of a rigid and stereotyped system.

It is to be hoped that a new generation of German academics will have a more open view of the role of women in the scientific community.

MARIA EUGENIA MONETA

Mechthildenstr. 19, 8000 München, FRG

1. Walton, A. *Chem. Brit.* May, 461–462 (1985).
2. Ruschhaupt, U. & Hartung, D. *MPG-Spiegel* April 22–26 (1988).

New telescope

SIR—Steven Dickman’s News item “Six feet off the ground” (*Nature* **341**, 177; 1989) gives an inaccurate representation of West German efforts to design a new large telescope (DGT).

Bochum and Göttingen are far from being rivals in the project. It is being pursued jointly by ten astronomical institutes in Germany including university departments, Max-Planck-Institutes and observatories. They cooperate within the DGT Working Group, of which the Göttingen and Bochum University institutes are also members. Alternatives for the telescope mounting such as the two referred to by Dickman are being studied in collaboration with the companies Krupp and MAN following a decision of the working group. The best solution will eventually be selected entirely on its technical and scientific merits. A Ritchie – Chrétien system based on a minimally segmented 12-metre primary containing a central light-weight 8-metre monolith is the optical configuration at present being explored by the workshop group with Zeiss. There is no race between Krupp and Zeiss: both companies are cooperating in this project. Together with MAN and DSD Dillinger Stahlbau, they are members of the DGT-Industries-Consortium.

The funding of the project is at present uncertain because West Germany is committed to making large contributions to the European Southern Observatory’s Very Large Telescope. Certainly there is not — and cannot be — a race between two small university groups such as Bochum and Göttingen to win a DM200 million project: it can be realized only by a national or — more likely — international effort.

KLAUS J. FRICKE

TH. SCHMIDT-KALER

*Universitäts-Sternwarte,
Geismarlandstrasse 11,
D-3400 Göttingen, FRG*

Eighteen ninety and all that

W. F. Bynum and J. L. Heilbron

Up for commemoration this year are anniversaries of the death of Benjamin Franklin and the births of H. J. Muller, August Möbius and (perhaps) Dom Pérignon. There is, too, the advent of the electric chair to 'celebrate'.

REGULAR readers of this anniversarial round-up will know that its authors aim not only to cull best bets in scientific anniversaries, but also to perfect studies of the anniversarial art. This year we introduce two new features — 'the anniversary of the year' and the concept of the 'censual'.

Our chosen celebrity for 1990 is the Anglo-American Benjamin Franklin. When he died two hundred years ago, there perished an entire academy of science, arts and belles lettres. He began his conquest of the republic of letters on the fringe of scientific low life, as the publisher of *Poor Richard's Almanac*, with a device that may still have applications. He announced in his first number that the market leader in the field of almanacs, one Titan Leeds, would soon and infallibly die — the stars said so, and Leeds himself concurred in the calculations, in so far as one astrologer could agree with another. The unhappy day came, as scheduled, in October 1733. Leeds's almanac continued to appear, giving sufficient proof that its author was dead. For, says Franklin, quoting from the latest edition of Leeds, "No Man living could or would write such Stuff". More is said about Franklin in his place, under 2.0 centenaries.

We have retained our usual format, beginning with centennials, continuing to sesquicentennials and so on, and ending with events that took place less than a hundred years ago. Here we have long felt an awkward want of terminology. For, as Adam knew, and Linnaeus confirmed, naming is close to knowing. Therefore, subject to market acceptance, we take as our term of reference the 'censual', symbol ϵ , which is the multiple by which 100 is turned into a sub- or super-centennial. This year Franklin's death has a censual of 2.0, that of Leeds, 2.57. A happier term is not likely to be found. It expresses at once the idea of a centennial and the pleasure felt by those who celebrate it.

1890 (1.0 centenary)

Not all Franklin's murders were literary. He was fond of turkeys, which he believed should have been the national symbol of the United States; electrocuting them, he said, much improved their taste. The technique was first practised on criminals in 1890, as a result of a market strategy reminiscent of the battle of the almanacs.

The Edison interests, whose business rested on direct current electricity, arranged for their rival, Westinghouse, which built on alternating current, to win the contract for the electric chair at Sing Sing. The demonstration that a.c. could kill did not save d.c.

Three other striking applications of electricity introduced in 1890 marked the

Ann Ronan

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Up, up and away — the kite experiment of Benjamin Franklin, innovator *extraordinaire*, which demonstrated the principle of the lightning conductor.

beginning of great industries. During his studies of nerve impulses, a French professor, Edouard Branly, invented a useful version of the 'coherer', a tube full of iron filings capable of detecting radio waves. Marconi began by improving Branly's device. An American engineer, Herman Hollerith (born 1860, 1.3 ϵ) developed a system for recording data on punched cards, to be read electrically, for the US census of 1890. It counted 63 million people according to several characteristics, each of which required the punching of a billion holes. Hollerith made his cards the size of dollar bills, to save on cutting equipment; the connection with money makes a good symbol, as Hollerith's patents were the foundation of the company that became International Business Machines (IBM). The third electrical invention of 1890, the 'teleautograph', was the brainchild of Elisha Gray, a founding partner of Western Electric and a fierce competitor of Alexander Graham Bell.

The teleautograph, which transmitted handwriting, found considerable use in banks and railroad stations. It was the original fax machine.

Science and commerce were, or became, intertwined in the newsworthy pronouncements of two German doyens of medical science. Robert Koch announced in August that he had at last discovered a substance capable of preventing the growth of the tubercle bacillus *in vitro* and *in vivo*. Koch's role in the original identification of the bacillus made the world beat a path to his doorstep. Tuberculin, his new product, had its uses, but they were more modest than the ones he first enumerated. In December, his sometime associate and longtime rival, Emil von Behring, burst upon the world with the idea, and some of the reality, of antitoxin, as both therapy and prophylactic. Tetanus and diphtheria were both within von Behring's ken, and although the important collaboration of his Japanese student, Shibasuro Kitasato, was forgotten by the time von Behring piped Koch to the post for the first Nobel Prize for Medicine or Physiology in 1901, it was diphtheria antitoxin — largely von Behring's brainchild — that was named as the reason for the award. The idea that blood or serum might be the mediators of important immunological defences was controversial in 1890. It went against the dominant cellular basis postulated by Rudolf Virchow (born 1820, 1.7 ϵ) and others. Nothing Virchow ever did had the commercial possibilities of the pronouncements of Koch and von Behring, however, and by the end of the century the German pharmaceutical industry was backing the new therapeutics in a big way.

The question of whether all bodies fall with the same acceleration in a space void of air played a part in the battle between the cosmologies of Descartes and Newton. The demonstration of the equal descent of the feather and the guinea dates back to around 1710 (2.8 ϵ). Precise measurement of the response of substances of different densities to the pull of gravity is more difficult than making a vacuum. In 1890, Roland Eötvös, professor of physics at the University of Budapest, harvested the first fruits of his torsion balance: all the substances weighed experienced the same gravitational attraction per unit mass. His later investigations with similar machines pointed towards big financial and scientific

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Westinghouse's a.c. electric chair, as depicted in *La Nature* (Paris) in September 1890. A laconic note pointed out that three surges of current were needed to finish off the criminal Kemmler, the first person to be executed in the chair.

gain. The balance detected irregularities in the intensity of local gravity and equality between the measures of gravitational and inertial mass. The first indicated where valuable minerals might be found; the second furnished one of the bases of general relativity (see under 0.5 ± 0.25 centenary).

Raising our gaze to the quintessence, we record the centennial of the first instalment of the *Draper Catalogue of Stellar Spectra* — it contained 10,351 stars down to the eighth magnitude. The work was done at Harvard University, under E. C. Pickering, with the support of a gift from the widow of the astronomer Henry Draper, and with the labour of a corps of women calculators under the command of Williamina Fleming. Among the lines on the Harvard plates were those that Pickering ascribed to hydrogen, and which, reassigned by Niels Bohr to helium in 1913, constituted decisive confirmation of the quantum theory of the nuclear atom. Thereafter the chief guides to the structure of atoms were the series spectra represented in the style introduced by Janne Rydberg, a lecturer in physics at the University of Lund, in his *Recherches sur la constitution des spectres*, of 1890.

Among the vital statistics of this year a century ago were the deaths of Christoph Hendrik Diederik Buys Ballot, who from his base in the Netherlands led in the organization of international meteorology, and garnered the eponym for the rule that he who stands with his back to the wind has a low atmospheric pressure on his left; James Nasmyth, a Scotsman turned Manchester entrepreneur, the inventor in around 1840 (1.5¢) of a useful steam hammer for shaping large forgings; Sir Edwin Chadwick, the English public health reformer who, despite a passion for the public good, "babbled too much, not of green fields, but of sewage"; and Heinrich Schliemann, born in Mecklenburgh-Schwerin, who began as a grocer's appren-

tice, made a fortune in indigo, dug up ancient cities, and gazed upon the face of Agamemnon.

Those born in 1890 include Sir William Henry Bragg, a Cambridge graduate, who began research at the age of 40, when professor of physics at the University of Adelaide, with such success that he was called to a chair at Leeds, and won a Nobel prize; Sir Ronald Aylmer Fisher, the myopic statistician of natural selection and much else, who, after professorships at University College London and Cambridge, emigrated to Adelaide; H. J. Muller, the American geneticist who taught us to value fruit flies; Kurt Lashley, the American neuropsychologist who learned much about vision but could not decide between the two halves of his brain; and Arthur Holmes, the English geophysicist whose pioneering work on radioactively determined geological time set the age of the Earth at over four and a half thousand million years (4.55×10^9).

1840 (1.5 centenary)

In the months ahead, energetic celebrants will want to commemorate the sesquicentennial of major moves towards thermodynamics. In July 1840, the physician Robert Julius Mayer deduced from the light colour of venous blood drawn from ailing sailors in Java that body heat and muscular motion are both derived from the "chemical force" inherent in food. Because less 'food force' was needed to keep warm in the tropics (hence the colour of the venous blood), more would be available for mechanical effort — good news for philosophers, bad news for malingering sailors. In December 1840, the brewer James Prescott Joule, working in Manchester where venous blood is dark, showed that it is not easy to keep warm with the heat generated by a current from an electric battery. He needed the most sensitive thermometers of the day to determine the rule now known as Joule's

law, that the heat developed by a current i flowing through a resistance R is $i^2 R$.

Mayer and Joule each continued in his own way, the philosophical doctor proceeding from lofty generalizations to occasional calculations, the brewer's son from delicate measurements to lofty generalizations. They arrived at the doctrine of energy conservation independently and roughly simultaneously, providing mutual reinforcement for a fundamental new principle and fuel for a heated dispute between the Germans and the British over whose man had priority of discovery.

High above the blood, the battery and the battle soared Nicolai Ivanovich Lobachevsky, correcting the proofs of his *Geometrische Untersuchungen zur Theorie der Parallellinien*, the most accessible exposition of his non-euclidean geometry. It impressed Gauss, who learned Russian to know more, but very few others before the 1860s. Justus von Liebig also wrote a book in 1840 that became famous. His *Die organische Chemie in ihrer Anwendung auf Agrikultur und Physiologie* concerned not the logic of parallels but the theory of dung. (British efficiency meant that the English translation was published first.) The book was initially more successful than Lobachevsky's, indicating that society needed manure more than mathematics. Liebig's recipes for supplying the nutritional requirements of plants fell out of favour at about the same time that mathematicians came to recognize the enduring value of Lobachevsky's work.

There was ice as well as fire in 1840. Louis Agassiz's *Etudes sur les glaciers* convinced the geological community that Joule's Manchester had seen even colder times. Glaciers were still common in Agassiz's native Switzerland, but he eventually settled in the New World, where in this year, 150 years ago, the Association of American Geologists was founded. Back in the Old, there died Heinrich Olbers, who thus (it is hoped) found the answer to his paradox, why the sky is dark at night; Johann Friedrich Blumenbach, whose doctoral thesis of 65 years previously is an anthropological classic; and Siméon-Denis Poisson, who held that life was good for only two things, to study mathematics and to teach it. As an offset there were born on 3 October the *Provincial Medical and Surgical Journal*, which grew up to become the *British Medical Journal*; Felix Anton Dohrn, the founder of the Zoological Station at Naples; and Ernst Abbe, who became the technical brain of the Zeiss optical company and, through his funding of the Carl Zeiss Foundation, a benefactor of scientific research.

1790 (2.0 centenary)

Speaking of foundations, in 1790 Friedrich Albrecht Carl Gren (born 1760, 2.3¢), professor of physics and medicine at the University of Halle, published the first

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Ancient art — one of the glories of the wall paintings at Lascaux, which were first revealed to modern eyes 50 years ago.

number of the *Journal der Physik*. Continued as the *Annalen der Physik*, it became for a time the leading physics journal; this year, *deo volente*, it will publish its 514th volume. Other important foundations of 1790 included the southernmost astronomical observatory in Europe, established in Palermo by Giuseppe Piazzi, and the US capital in Washington, DC, which dispenses every ten minutes more money for science and technology than the entire world spent on them annually two centuries ago.

In 1890 the president of the American Philosophical Society (founded 1740, 2.5¢) observed that whereas a single individual could deliver a necrology of Franklin at the time of his death, it took no fewer than five experts to celebrate him a century later, so consequential and variegated had his contributions proved. The expert of 1890, charged with itemizing Franklin's contributions to science and technology, classified them as follows: health and sanitation (smoke abatement, clean water supply, ventilation of public buildings, domestic heating, bifocals); navigation (study of the Gulf Stream, oiling of waters); meteorology (development of storms, origins of the aurora); and electricity (conception of positive and negative electricity, invention of the lightning rod). Franklin proposed the experiment of drawing lightning to a pointed conductor in 1750 (2.4¢); it was first tried in France, and succeeded in detecting electrical disturbances in the lower atmosphere. The first electrician to catch a thunderbolt was Georg Wilhelm Richmann. It killed him. "Fear not death", sayeth Poor Richard; "for the sooner we die, the longer shall we be immortal."

Scientists born in 1790 include John Herapath, an English journalist who prematurely revised the kinetic theory of

gases; August Ferdinand Möbius, a German mathematician responsible for the mind-tingling 'Möbius strip'; and Claude-Servais-Mathias Pouillet, a French physicist best known for his measurements on gases and electrical resistance. They partly compensated for the death of the Scottish chemist and physician William Cullen, who had been father figure as well as teacher to generations of Edinburgh medical students.

1740 (2.5 centenary)

In this year, Titan Leeds invented a novel way to communicate with the living: "Finding you [Poor Richard] asleep, I entered your left Nostril, ascended into your Brain, found out where the Ends of those Nerves were fastened that move your right Hand and Fingers, by the help of which I am now writing unknown to you. . .". As Leeds ended his inventive career, that of Michel Joseph Montgolfier began; 43 years later, he and his younger brother Etienne Jacques flew a hot-air balloon, the first to carry human passengers. Two other pioneers of the study of the air were born in 1740: Louis Cotte, a great organizer and propagandist for meteorology, in which he included earthquakes, terrestrial magnetism and the aurora borealis; and Horace Bénédic de Sausurre, who helped perfect the barometer, thermometer and hygrometer, and took them with him on his many Alpine ascents, to measure the state of the atmosphere as close to Heaven as possible.

Those who like to celebrate quarter millennia may wish for more, and less usual, items to choose from. Know then that in 1740 Abraham Trembley, born in 1710 (2.8¢), began to multiply hydra by cutting them to pieces. Growing a full animal from a tentacle created a sensation. So did William Stukeley's *Stonehenge*,

which juxtaposed realistic accounts of excellent field work at the monument with wild speculations about its use by the Druids as a centre of the rite of the "reform'd patriarchal Religion" of Abraham. The book is further enriched by engravings of Druids, one with the likeness of Stukeley. While Stukeley updated the religion of Britain, the Church of Rome updated itself, by electing Prospero Lambertini as Pope. As Benedict XIV, he continued his benefactions to science, particularly in his home town, Bologna, then a papal possession. He gave its academy of sciences a collection of the best contemporary apparatus for experimental physics; he commissioned the creation of eight life-sized wax figures for its department of anatomy; and he made possible its acquisition of a most important set of obstetric models to help ensure the safe delivery of new souls.

1690 (3.0 centenary)

In or around 1690, Dom Pérignon, cellarer to a Benedictine abbey near Epinay, may or may not have invented champagne. He certainly much improved it. Some say he was born in 1640 (3.5¢) and so qualifies for notice by birth; but no one will object to celebrating the third centennial of the anniversarial beverage *par excellence* on the fussy grounds that we do not know when it was invented. To help with the morning after, John Locke published his *Essay Concerning Human Understanding*, exactly two centuries before sober William James contributed his own *Principles of Psychology*.

One of the grandest of all exercises in human understanding was published in 1690, in Christian Huygens's *Traité de la lumière*. It threw light on light (to use a favourite phrase from Franklin) by presenting a consistent wave model, and gave a construction and explanation of the properties of birefringent crystals. It offered a base for battle against the particulate theory of light until 1810 (1.8¢), when another of this year's honourees, S.-D. Poisson (see under 1840), explained the principal phenomena of double refraction and polarization as consequences of short-distance forces acting on oblong particles of light, an illusory success soon undercut by Huygens's approach, as revised by Thomas Young and Augustin Fresnel.

1590 (4.0 centenary)

For those able to care, in 1590 Tycho Brahe began the series of observations that led him to the discovery of the third inequality in the motion of the Moon. Some anniversarial guides give us leave to date the microscope (and/or the telescope) to 1590, to the precocious optician and counterfeiter Zacharias Jansen. Although Jansen was only two years old in 1590, his son, Sacheriassen, testified in

1655 that his father had presented the scientific world with an important new instrument then. It is symptomatic of the pitfalls of anniversary studies that the pious claim was believed by not a few people afterwards. The year saw the deaths of Giovanni Batista Benedetti, born 1530 (4.6¢), a mathematician whose writings on motion and mechanics influenced Galileo; and the surgeon Ambroise Paré, born around 1510 (4.8¢), who, by dressing gunshot wounds while letting God cure them, revolutionized (for the better) their treatment.

1540 (4.5 centenary)

This was a year full of good books and other novelties. Peter Apian published his beautiful *Astronomicum caesereum*, full of comet tails and volvelles, arranged in the geocentric manner, just as Georg Joachim Rheticus brought out his *Narratio prima*, containing the first report of Copernicus's views about the structure of the Universe. Vannoccio Biringuccio (born 1480, 5.1¢) gave to the world his *Pirrotechnia*, an underground book in two senses: it treats mining and metallurgy, and it was issued posthumously. Garcíá López de Cárdenas became the first European to gaze at the Grand Canyon; the souvenir shops did not come until later. Among those who entered this life in 1540 was François Viète. His most important book on algebra, and the earliest sustained work on symbolic algebra, *In artem analyticam isagoge*, was published in 1591 — rating a 4.0¢ to within a quarter of a per cent.

1190 (8.0 centenary)

This year's outliers are Vincent of Beauvais, a Dominican encyclopaedist, author of the *Speculum maius*, or great mirror of all learning, which was completed with funds from the treasury of Louis IX of France; and Villard de Honnecourt, the builder and architect, whose notebooks give unique access to the technique and self-image of a mediaeval master craftsman. Both were born in 1190, approximately.

1940 (0.5 centenary)

War buffs will find much of note in the semi-centennials of the invention of the cavity magnetron, the heart of centimetric radar; of the visit to the United States of a technical group headed by Henry Tizard, which brought the magnetron and other new implements of war to the attention of American military and scientific men; of the foundations of the Radiation Laboratory at the Massachusetts Institute of Technology to develop radar and of the National Defense Research Committee to mobilize American scientists for war work; and of experiments that identified uranium-235 as the fissionable ingredient of natural uranium and increased the likelihood that physicists would be able to

release the energy of the nucleus.

In 1940, The Rockefeller Foundation gave more than a million dollars to the University of California towards the construction of a huge cyclotron, intended as an affirmation of human rationality and of the disinterested pursuit of knowledge in an increasingly insane world; but the machine soon participated in the insanity by helping to perfect the means to make uranium bombs. Meanwhile, still in 1940 and still at Berkeley, Edwin McMillan and Philip Abelson detected the first bona fide transuranic element, neptunium, an essential step towards the recognition of another nuclear explosive, plutonium; and Martin Kamen and Samuel Rubin discovered carbon-14, an important tracer for biochemical reactions and radioactive dating.

"It looks like a miracle", Howard Florey supposedly told his wife-to-be, when the results of his 25 May experiment with mice, streptococci and penicillin were known. If this seems out of character for the formidable professor of pathology at Oxford, since then the word 'miracle' has often been attached to the drug he and his group developed. Many birthdays could be claimed for penicillin, but it has had a thriving existence since 1940.

Some infants are also more likely to thrive because of another discovery in that first full year of war — the identification by Karl Landsteiner, Philip Levine and Alexander Wiener of the Rhesus (Rh) factor in human blood. It helped to explain why some blood transfusions, though compatible on the major system of blood groups (A, B, AB, O) elaborated much earlier by Landsteiner, produced serious reactions. Levine extended the work to fetuses, which, inheriting the positive Rhesus factor from their fathers, could provoke their mothers to manufacture antibodies against their own blood cells.

No Nobel prizes were awarded in 1940; Albert Einstein then (but not therefore) swore allegiance as a citizen of the United States; and Henri Edouard Prosper Breuil, an ordained priest and professor of prehistory at the Collège de France, became the first archaeologist to see the splendid wall paintings in the caves at Lascaux. Among casualties unrelated to the war were Carl Bosch, co-inventor of the process for drawing nitrogen from the air; Heinrich Johannes Gustav Kayser, insatiable measurer of spectral lines; Oliver Lodge, electrician and spiritualist; Dmitry Sergeevich Rozdestvensky, founding director of State Optical Institute in Leningrad; Julius Wagner von Jauregg, who believed malaria was preferable to tertiary syphilis; Hans Zinsser, who made the louse into a kind of art form; Joseph John Thomson, discoverer of the electron; and Pierre Weiss, non-discoverer of the magnetron.

1915 and 1965 (0.5 ± 0.25 centenary)

The news of 1915 confronts the anniversary with the best and worst in human history. On the down side, the acceleration of the perversion of science and technology to military purposes — the first use of poison gas; on the up side, the completion of that monument to pure thought, the general theory of relativity. Further to the negatives was the death in battle of many prominent and promising scientists on both sides; H. G. J. Moseley may stand, or rather fall, for them all. Further to the positives were Arnold Sommerfeld's discovery and derivation of the fine-structure constant, which first combined relativity theory with the quantized atom; Joseph Goldberger's brilliant epidemiological studies on pellagra; Pyrex glass; and the first transcontinental telephone (New York to San Francisco) and the first radio-telephone conversation (Paris to Arlington, Virginia). The call to San Francisco opened an International Exposition, in which the neutral United States celebrated the ploughs and pruning hooks that the Europeans had made into swords and spears. The publication of Alfred Wegener's *Die Entstehung der Kontinente und Ozeane* might have alerted prescient readers that Europe's problems could not be forever isolated, since the Old World and the New had originally been One.

There is much to commemorate about 1965. We had the discovery by Arno Penzias and Robert Wilson of the background radiation interpreted as a memorial of the Big Bang; the smaller bang marking the start-up of the Stanford Linear Accelerator; the detection at the Arecibo Observatory that Venus rotates backwards and that Mercury does not (as reported here last year, under 1889) keep one face permanently towards the Sun; the detection in a South African gold mine, by Frederick Reines and J. P. F. Sellshop, of the elusive neutrino; soft contact lenses; and a variety of space walks by American and Soviet astronauts.

Sometime between the Big Bang and the space walks Stonehenge came to be. In 1965, the astronomer Gerald Hawkins gave, in his *Stonehenge Decoded*, an interpretation more in keeping with our technical age than Stukeley's romance about Druids. The astronomer's point of departure is even less plausible than the antiquary's: "If I can see any alignment, general relationship or use for the various parts of Stonehenge, then these facts were also known to the builders". Next year we may report on the application of this starting principle to the anniversary method.

W. F. Bynum is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BP, UK. J. L. Heilbron is in the Office of History of Science and Technology, University of California, Berkeley, California 94720, USA.

Aren't we all green now?

By neglect of consultation and expertise, the British government's new pollution bill seems designed to make even sensible decisions on environmental pollution and genetic engineering hostage to public anxiety and suspicion.

CRITICS of the British government's green bill, published two weeks ago, have quickly labelled it a "hotchpotch" of unrelated proposals (see page 3), but that was always on the cards. In particular, it says nothing to reflect the government's more recent conversion to the belief that people (and governments) should worry about the excess greenhouse effect. In most respects, the bill is a good-house-keeping measure, intended to tie up loose ends left dangling from administrative decisions in the past several years.

Many of the inspectors who will be members of the grandly named Her Majesty's Inspectorate of Pollution, for example, are already at work at the Department of the Environment, doing what they can when there are few of them to practise the new doctrine of Integrated Pollution Control. Similarly, there is nothing new in the decision that the British Nature Conservancy Council should be split into three parts geographically and then united with the old Countryside Commissions for Scotland, Wales and England; all that was announced last summer by Mr Nicholas Ridley, then the Secretary of State for the Environment.

But none of this implies that the bill is a formality, or that it should be treated as such when it comes to be debated in the British parliament. On the contrary, there are several issues that deserve urgent attention, not the least of which is the fate of the Nature Conservancy Council (NCC).

Briefly, this neatly illustrates how governments can remove thorns in their flesh by the process of emasculation. Forty years ago, NCC was what would be called a mission-oriented research council. The mission was the conservation of British flora and fauna and of tracts of natural environment supporting them. The council had only limited statutory powers, being required to rely on government ministries for legislation to serve its aims, but it supported a substantial research effort through laboratories of its own and grew to be an influential voice on several public issues — the use of pesticides in agriculture, for example, and proposals to flood valleys so as to make reservoirs, on which it often intervened as an objector in the public inquiries required by British planning law. To British governments (not just the present government), NCC has been a particular nuisance because of its power to designate certain sites as having Special Scientific

Interest, which impedes development and sometimes requires the payment of substantial sums in compensation.

Constitutionally, the separate existence of NCC (until now) must be supposed to mark an earlier enlightened recognition that there would always be a potential conflict between the interests of governments (which include development) and of conservation (which tends to be inimical thereto). In much this spirit did earlier British governments set up the University Grants Committee to remove the distribution of public funds from day-to-day political considerations.

As events have shown, governments and their civil servants are not easily content with the public support of potential adversaries. The first major step in the emasculation of NCC was the removal of its research arm (as the Institute of Terrestrial Ecology) to the Natural Environment Research Council, on the grounds that the council could always commission research that seemed particularly urgent. Now it, together with its small army of working naturalists, is to be dismembered. The result will be a further attenuation of the scientific influence on British conservation policy. While the civil servants will find that rows over conservation policy do not abruptly cease, they can but hope that conservationists will in future make a less, or even less, coherent case. Will outright abolition of institutionalized conservation then be practicable?

The new proposals for the regulation of the release of genetically engineered organisms raise related questions. That regulations of some kind are necessary is generally accepted, and will in any case be required of Britain by the European Community. By now, as it happens, there is a thoroughly tested way of dealing with problems such as these in the related field of the manipulation of organisms in the laboratory and in manufacturing industry. People send proposals to a committee, together with a hazard assessment, are given or denied consent to go ahead (policed, in Britain, by inspectors of the Health and Safety Executive) and a report is eventually published from which people can learn what genetic manipulation goes on within their midst. Cumbersome though it may be, the device has helped enormously in quietening public anxiety about genetic manipulation.

The most serious defect of the new green bill, in its proposals on the release of

engineered organisms, is that it turns its back on this valuable example. The bill says nothing about the expert committee the Department of the Environment has separately promised will vet proposals for the release of engineered organisms. Particular proposals, the bill says, will be made by the Secretary of State for the Environment, who will also decide what conditions should be attached to any consent he may give for the release of genetically engineered organisms, which conditions may or may not be published.

Such an arrangement seems calculated to excite disquiet. A committee, however expert, lacking statutory authority will carry less weight than one whose scrutiny of proposals is required by law, while the ambiguity about the government's intentions on the disclosure of the terms on which it will allow field experiments is a recipe for creating distrust. Those with plans for sowing newly engineered varieties of plants in British fields, who may be relieved that the British government has decided not to follow the over-restrictive proposals of the Royal Commission on Environmental Pollution, should be alarmed that opportunities for public alarm are being so needlessly excited. Even in the British constitution, there are precedents: the Committee on the Safety of Medicines has statutory existence.

The other lacuna in the British green bill is potentially even more serious, but the grand concept of what is called Integrated Pollution Control is not even defined. Innocents may conclude that the people previously known as the Alkali Inspectors have merely had the scope of their responsibilities broadened so as to cover discharges of all sorts from all kinds of undertaking, not only private but public.

But the details of how this will be done are again left for the secretary of state to decide, and will be issued as separate regulations. The bill does not even require the minister to take account of, say, the absorptive capacity of a watercourse for pollution when promulgating regulations and, while the objective of the whole exercise is defined as the prevention of 'harm' to people and other living things, there are no criteria for deciding how degrees of harm of different kinds should be balanced against each other. Again the government is assuming that it (and its successors) will, by definition, know best, which is hardly the way in which to command public confidence. **John Maddox**

The club-sandwich mystery

J.D. Mollon

"THE whole organization of the human lateral geniculate nucleus (LGN) cries aloud that something is being segregated. The question before us here is simply: *what?*" It was with this question that the late Gordon Walls opened his engaging monograph on the LGN, the knee-shaped nucleus of the mammalian brain that lies in the visual pathway between the retina and the striate cortex¹. A paper by Schiller, Logothetis and Charles on page 68 of this issue² offers an answer to Walls's question.

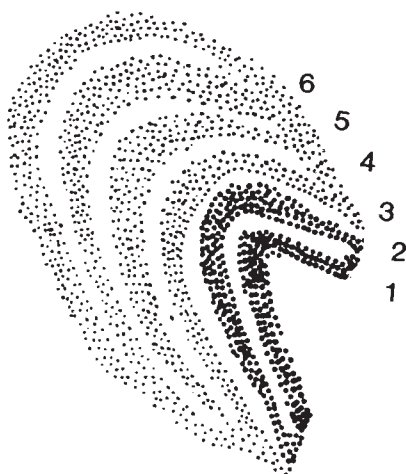
Medical students are still told that the LGN is a mere relay station. But to the neurobiologist the LGN is provocative in its exquisite organization. In Old World primates and in man, the lateral geniculate consists of six layers, each eye providing the inputs to three layers (see figure). The two layers that lie on the inside of the geniculate 'knee' consist of large cells and hence are known as the 'magnocellular laminae'; their main retinal inputs are from the axons of the large cells that Perry and Cowey³ called P α cells. The remaining four layers consist of small cells and so are known as the 'parvocellular laminae'; their main retinal inputs are from the class of cells known as P β cells. If we number the six laminae from the inside of the knee outwards, then layers 1, 4 and 6 of each geniculate are drawn from the contralateral eye and layers 2, 3 and 5 from the ipsilateral. Each lamina forms a map of the contralateral half of the visual field, and, what is most remarkable, the six maps are in precise alignment⁴. Gordon Walls¹ likened the lateral geniculate to a club sandwich: the toothpick piercing the sandwich corresponds to a single direction in visual space.

Why is the lateral geniculate stratified? How is the task of analysing the visual world distributed among the laminae (and their cortical projections)? Why does the developing visual system go to such trouble to align the six maps in the LGN? And why indeed does the LGN exist? These must be interdependent questions. But it is the second of them that has been most explicitly addressed, and three types of evidence have been offered.

First, electrophysiological recordings have been made from the primate LGN while various stimuli were presented to the eye. When the stimulus varies only slowly in space (that is, it is of low spatial frequency), parvocellular units typically exhibit chromatic selectivity, being excited by one part of the visible spectrum and inhibited by another part. Because the opposing receptor inputs to such cells

are arranged in concentric areas of excitation and inhibition, and because these 'receptive fields' are small, parvocellular units are also sensitive to stimuli of high spatial frequency, whatever the stimulus wavelength. By contrast, the magnocellular units tend to respond maximally at lower spatial frequencies, exhibit little chromatic selectivity and are often more transient in their response⁵⁻⁷.

One might think that the functions of the different LGN layers would be readily clear from such studies of single cells. But in fact there has been fierce



A cross-section of the lateral geniculate nucleus of a typical Old World primate, showing the conventional numbering of the magnocellular (1–2) and parvocellular (3–6) laminae. Each of the six laminae contains a separate map of the contralateral visual field, the central region of the field being represented in the central part of each part lamina. The segregation of visual pathways, which becomes so patent in the LGN, is now known to begin in the outer plexiform layer of the retina and to continue beyond the LGN in the striate and prestriate cortex.

controversy about the role of the parvocellular system in the analysis of spatial patterns. One extreme view was adopted by Shapley and Perry⁵, who pointed out that magnocellular units are much more sensitive to spatial contrast than are parvocellular units and thus may be responsible for the spatial analysis of low-contrast stimuli even at relatively high spatial frequencies. This account, in which colour vision becomes the primary office of the parvocellular system, is rather implausible — there are about 7.5 times as many parvocellular units as magnocellular⁸; the magnocellular system does not sample the retina with the resolution needed to account for psychophysical performance⁹; and there is a well-developed parvocellular system in dichromatic platyrrhine monkeys. (For the detailed

arguments see refs 5–7 and 9.)

Why has the electrophysiological approach been so inconclusive? One reason is that we cannot properly compare the discriminatory performance of individual cells in layers 1 and 2 with the performance of individual cells of the parvocellular layers. If the magnocellular system is the more primitive or less encephalized system ("protopathic" in Henry Head's terms), then it is likely to have carried out much of its spatial integration before the LGN, whereas the "epicritic" parvocellular system is more encephalized and postpones its spatial integration to the cortex. So it is unjust and artificial to use a cell-for-cell comparison to judge the two sets of signals as they hurry through the LGN.

A second approach has been psychophysical: here the experimenter presents, to a human observer, stimuli that are thought to excite only one of the two channels. Particular favourites have been 'equiluminant' stimuli (stimuli that vary only in colour and not in luminance), which are held to silence all but the parvocellular pathway. Several perceptual functions (such as stereopsis, movement discrimination and figure-ground differentiation) are impaired when the stimuli are equiluminant, whereas others survive⁷. It is concluded that the signals of the colour-selective laminae are unable to sustain those perceptual functions that fail at equiluminance, but are able to sustain those that survive. But these arguments depend fatally on a reification of the psychophysical construct of 'luminance'¹⁰.

Consider first a perceptual function that does survive when the stimulus is a red figure on an equiluminant blue ground. At the edge between figure and ground the long-wave cones must detect a spatial transient of one sign and the middle-wave cones must detect a transient of the opposite sign. We are asked to believe that the two transients will exactly cancel each other out in all the cells of the magnocellular pathway, a pathway that seems especially designed to detect transients. This is contrary to biological plausibility¹⁰ and to empirical evidence¹¹. Consider now a perceptual function that is impaired at equiluminance. We need not conclude that this task is primarily the charge of the magnocellular system. For equiluminance presents a potentially disruptive stimulus to the parvocellular system: whereas long-wave on-centre units of the P β type normally give the same sign of response to spatial transients as do middle-wave on-centre units, this yoking is lost at 'equiluminance', when the two types of on-centre unit (and the two types of off-centre unit) will give responses of contradictory sign¹².

The third approach is to make selective lesions in the LGN. Merigan and Eskin¹³

tested monkeys after systemic administration of a neurotoxin, acrylamide monomer, which damages the small-cell system but seems to spare other pathways: the monkeys were still well able to detect high temporal frequencies or low spatial frequencies, but were badly impaired at low temporal frequencies and high spatial frequencies. More recently, the same laboratory has reported the consequence of making a local LGN lesion by injecting ibotenic acid into the magnocellular layers: the impairment was then apparent only at high temporal frequencies and low spatial frequency¹⁴.

It is lesion experiments of this type that are now reported by Schiller, Logothetis and Charles. They have trained their monkeys to deflect their gaze to the part of the visual field that contains a target stimulus and have then made local lesions in the LGN. Their answer to Wall's question is one that would have been conventional in the mid-1970s, before the recent controversies and before the mischievous of equiluminance was abroad. The

parvocellular laminae appear to carry the signals that mainly sustain perception of colour, texture, shape and fine stereopsis, whereas the magnocellular signals sustain the detection of movement and flicker. □

J.D. Mollon is in the Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

1. Walls, G.L. *The Lateral Geniculate Nucleus and Visual Histophysiology* (University of California Press, Berkeley, 1953).
2. Schiller, P.H., Logothetis, N.K. & Charles, E.R. *Nature* **343**, 68–70 (1990).
3. Perry, V.H. & Cowey, A. *Expl Brain Res.* **43**, 226 (1981).
4. Kaas, J.H., Guillery, R.W. & Allman, J.M. *Brain Behav. Evol.* **6**, 253 (1972).
5. Shapley, R. & Perry, V.H. *Trends Neurosci.* **9**, 229 (1986).
6. Lennie, P. & D'Zmura, M. *CRC Crit. Rev. Neurobiol.* **3**, 333 (1988).
7. Livingstone, M. & Hubel, D. *Science* **240**, 740 (1988).
8. Le Gros Clark, W.E. *J. Anat.* **75**, 419 (1941).
9. Mollon, J.D. & Jordan, G. *Die Farbe* (in the press).
10. Mollon, J.D. *A. Rev. Psychol.* **33**, 41 (1982).
11. Lee, B.B., Martin, P.R. & Valberg, A. *J. Neurosci.* **9**, 1433 (1989).
12. Mollon, J.D. *Die Farbe* (in the press).
13. Merigan, W.H. & Eskin, T.A. *Vision Res.* **26**, 1751 (1986).
14. Merigan, W.H., Byrne, C.E. & Maunsell, J. *Soc. Neurosci. Abs.* **15**, 1256 (1989).

SOLAR SYSTEM

Source of short-period comets

Glen R. Stewart

BEYOND the orbits of Neptune and Pluto lies a vast region of space, extending 1.5 light years, that is populated with perhaps as many as 10^{13} comets. This region of the Solar System, commonly referred to as the Oort cloud, is postulated as the ultimate source of all known comets. Although it is not yet possible to observe such distant objects directly, much can be deduced about the Oort cloud from the orbits of the few observable comets that pass through the inner Solar System. In this context, the 100 known comets with orbital periods less than 15 years present an intriguing puzzle. Their orbits mostly lie close to the mean orbital plane of the planets, whereas the 500 known long-period comets follow orbits with inclinations distributed isotropically about the mean orbital plane. Where do short-period comets come from? Do long-period comets naturally evolve into low-inclination short-period orbits, or do the two groups of comets originate from different regions of the Oort cloud? Two recent theoretical studies reach opposite conclusions. Duncan, Quinn and Tremaine¹ conclude that short-period comets come from a low-inclination comet belt lying a short distance beyond Neptune's orbit. And in the 1 December issue of the *Monthly Notices of the Royal Astronomical Society*, Stagg and Bailey², criticizing that work, argue for a common source in an isotropic Oort comet cloud. Together, these studies provide tantalizing new clues about the structure and evolution of the outer

regions of the Solar System.

Both investigations attempt to simulate the long-term orbital evolution of numerous sample comets and therefore each employs an approximation to complete the calculation in a reasonable amount of computing time. Duncan *et al.* shorten the required integration time by artificially increasing the masses of the four giant gas planets — Jupiter, Saturn, Uranus and Neptune — by a factor of 40. This procedure does not change the statistical properties of the final distribution of comet orbits if successive gravitational perturbations by the giant planets lead to a gradual diffusion of a comet's orbit parameters. The authors tested the validity of their approximation by running some of their simulations twice, using a different mass scaling factor for each run. Nevertheless, Stagg and Bailey raise serious doubts about the accuracy of this assumption and describe calculations of comet orbits that evolve in large discrete steps owing to rare close encounters with the giant gas planets.

Planetary perturbations

The Monte Carlo computer simulation technique used by Stagg and Bailey replaces the numerical integration of perturbed comet orbits with a series of randomly selected planetary perturbations. The perturbations are drawn from a parameterized distribution that depends on the comet's semi-major axis, inclination and perihelion distance. Similar

Monte Carlo calculations were previously reported by Everhart³, who assumed that only one of the giant planets could perturb a comet orbit with a given set of orbit elements.

Stagg and Bailey generalize Everhart's procedure by allowing for perturbations from all four giant planets to act on any given comet. Because the probability of a long-period comet being perturbed into a short-period comet is quite small, Stagg and Bailey use a 'cloning' procedure to increase the number of comets that finally evolve into short-period orbits. In their original sample of 5,000 comets only about one in a hundred evolve inward to a semi-major axis less than 100 astronomical units. This subset of 57 comets is expanded to four times their number, or 228, before continuing their orbital evolution. The number of these comets that evolve inward to the orbit of Uranus are again multiplied to three times their number. Similarly, those that reach Saturn's orbit are multiplied twenty-five times. In this way Stagg and Bailey are able to calculate the probability of extremely rare events.

Close encounters

Perhaps the most surprising result of Stagg and Bailey's work is that extremely rare, close planetary encounters may actually dominate the long-term evolution of a comet's orbit. The source of this new result seems to be the inclusion of perturbations from all four giant planets, which tends to increase the probability of close encounters. Rather than gradually diffusing inwards from beyond the orbit of Neptune to an orbit that crosses Jupiter's orbit in many small steps, about a third of the comets in the Monte Carlo simulation jump from Neptune to Saturn in a single step. The evolution into a short-period orbit is much faster than predicted by diffusion, but the process is also less efficient so that the overall rate of capture of short-period comets found by Stagg and Bailey is only about twice the value found by Everhart.

It may seem paradoxical that the direct orbit integrations done by Duncan *et al.* would have missed the effects of rare close encounters because they increased the masses of the perturbing planets. The key property of these events is their rarity; by increasing the average size of the more common planetary perturbations, Duncan *et al.* may have caused the comets in their simulation to evolve into short-period orbits before they had a chance to experience a really rare close encounter. This tentative conclusion based on the Monte Carlo simulations obviously needs to be tested against direct orbit integrations.

Partly in response to Stagg and Bailey's criticisms, Duncan *et al.* have recently repeated their orbit integrations with a planetary mass scaling factor four-times

smaller than was used in their published work. Duncan *et al.* still find no evidence for the rare close encounters reported by Stagg and Bailey. It would be desirable to remove the planetary mass scaling completely, but the required computing time is prohibitive at present. To reduce the mass scaling factor from 40 to 10, the computing time had to be increased by a factor of 16. To remove the mass scaling entirely would require an integration time 1,600-times longer (S. Tremaine, personal communication).

Kuiper belt

On the basis of the results of their orbit integrations, Duncan *et al.* conclude that the source of short-period comets cannot be the isotropic Oort cloud's distribution of nearly parabolic orbits thought to give rise to the known long-period comets. Such a source always produces many more high-inclination short-period comets than are currently observed. Good agreement with the observed distribution of short-period comets was obtained when they postulated a source of low-inclination (less than 30°) orbits with perihelia near Neptune's orbit.

If confirmed, this result would provide strong evidence for the previously postulated Kuiper belt of comets extending from Neptune's orbit out to a distance of several hundred astronomical units⁴. Some current models of the origin of the Solar System produce a Kuiper belt of icy bodies in the process of assembling the giant gas planets. The recent detection of disks of debris of this size around several stars lends some credibility to the existence of a Kuiper disk⁵. However, these observations only prove the existence of micrometre-size dust particles; the existence of an associated comet belt can only be inferred.

Stagg and Bailey point out that a separate source of short-period comets in the Kuiper belt would not be necessary if planetary perturbations were less effective at changing the semi-major axes of high-inclination orbits. Duncan *et al.* found some dependence of this kind, but their use of supermassed planets may have obscured the true dependence on comet orbit inclination. Because Stagg and Bailey assumed that orbit inclinations were unaffected by planetary perturbations in their calculation, they could not test this idea either. Future Monte Carlo calculations that allow for changes in orbit inclinations should help resolve the issue. Nevertheless, Stagg and Bailey may have a difficult time trying to convince their sceptical rivals because the use of Monte Carlo methods to simulate very rare events is often a subtle and treacherous business.

Stagg and Bailey further speculate that the orbit distribution of short-period comets may be more isotropic than the

observations seem to imply. They find that the time required to capture a comet into a short-period orbit strongly increases with the assumed inclination of the orbit. For inclinations greater than about 30°, the capture time is likely to exceed the active lifetime of the comet. Short-period comets in high-inclination orbits would therefore be mislabelled as asteroids if they have already exhausted their supply of volatiles by the time they reach a short-period orbit. Wetherill has argued that the ten or so Apollo–Amor asteroids known to be in high-inclination orbits may well be extinct comet nuclei⁶. The discovery of many more such objects could mitigate the need for a separate source for short-period comets and would weaken the

evidence for the Kuiper belt beyond Neptune. More surprises may still await those who search for the source of comets. □

Glen R. Stewart is in the Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, Colorado 80309, USA.

1. Duncan, M., Quinn, T. & Tremaine, S. *Astrophys. J.* **328**, L69–L73 (1988).
2. Stagg, C. R. & Bailey, M. E. *Mon. Not. R. astr. Soc.* **241**, 507–543 (1989).
3. Everhart, E. in *IAU Coll. No. 39, Comets–Asteroids–Meteorites* (ed. Delsemme, A. H.) 99 (University of Toledo, Ohio, 1977).
4. Fernandez, J. A. *Mon. Not. R. astr. Soc.* **192**, 481–491 (1980).
5. Backman, D. E. & Gillett, F. C. in *Cool Stars, Stellar Systems, and the Sun* (eds Linsky, J. L. & Stencel, R. E.) 340 (Springer, Berlin, 1987).
6. Wetherill, G. W. *Icarus* **76**, 1–18 (1988).

NUCLEIC ACIDS

Adding to the genetic alphabet

Leslie E. Orgel

WHY are all nucleic acids assembled from just four kinds of monomer? Is it because there are only two kinds of base pair that fit into the Watson–Crick double helix? Or because evolution failed to 'discover' further base pairs or 'decided' that four bases are enough? In a paper from Steven Benner's group on page 33 of this issue, Piccirilli *et al.*¹ show that the second answer is probably the correct one. They have synthesized a nucleoside analogue κ that forms with xanthosine (χ) a novel base pair whose structure is compatible with the double helix. Neither κ nor χ pairs efficiently with any of the four standard nucleosides, so faithful copying of polymers containing all six bases by DNA and RNA polymerases is, in principle, possible. It has been shown that the pyrimidine analogue κ , when incorporated into an oligodeoxynucleotide, leads to the faithful enzymatic incorporation of χ into complementary DNA or RNA.

A strong base pair is joined by three

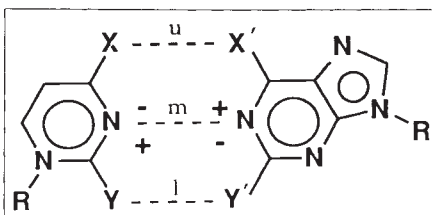


FIG. 1 An idealized strong base pair. If X is a donor NH₂ group, X' is an acceptor oxygen atom, and vice versa. The same rule applies to Y and Y'. The middle hydrogen bond (m) involves a single hydrogen atom, which may be present either on the pyrimidine or on the purine. For simplicity of representation I have supposed that the distribution of heterocyclic nitrogen atoms is the same as in a standard base pair. In fact, certain of the new base pairs can only be realized with altered arrangements of heterocyclic nitrogen atoms.

hydrogen bonds. In Fig. 1 the bonds are labelled 'u' (upper), 'm' (middle) and 'l' (lower). The middle hydrogen bond always joins two heterocyclic nitrogen atoms, one and only one of which carries a hydrogen atom. Each of the upper and lower hydrogen bonds joins an exocyclic donor amino group on one base to an exocyclic acceptor oxygen atom on the other. Thus there are, formally, eight distributions of hydrogen donors and acceptors that could work. The pyrimidine could carry either an oxygen or a nitrogen at the upper and at the lower exocyclic positions, and might or might not have a hydrogen atom on the central ring nitrogen. Once the structure of the pyrimidine is decided, the donor–acceptor structure of the purine that pairs with it is completely determined. Provided the bases cannot undergo tautomeric rearrangements, no strong interaction between inappropriate bases to give a strong base pair of standard structure is possible, because any 'cross-pair' would involve at least one potential hydrogen bond with zero or two hydrogen atoms.

Benner and his colleagues show that pairs of plausibly stable nucleotide analogues corresponding to at least six of the eight types of strong base pair can be formulated. In the familiar G:C base pair, cytosine has an NH₂ atom in the upper position, an oxygen atom in the lower position and does not carry a hydrogen atom on the central nitrogen atom. The A:U base pair is weaker because it makes use of only two of the three possible hydrogen bonds; 2,6-diaminopurine nucleoside, with amino groups at both exocyclic positions and lacking a hydrogen on the central nitrogen, forms three hydrogen bonds with uridine. Benner's group has studied two novel hydrogen-bonding schemes, one represented by the

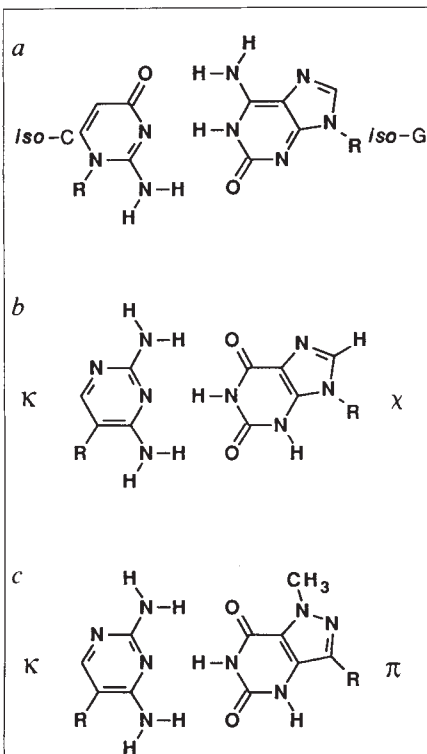


FIG. 2 Structures of some non-standard base pairs. *a* *iso-C:iso-G*. *b* $\kappa:\chi$. *c* $\kappa:\pi$.

iso-C:iso-G base pair (Fig. 2*a*) and the other the base pairs formed by κ with χ (Fig. 2*b*) or with an equivalent purine nucleoside analogue π (Fig. 2*c*) that is more readily incorporated into oligodeoxynucleotides by non-enzymatic synthesis.

The idea of a third base pair involving *iso-G* and *iso-C* is not entirely new². Benner and his co-workers previously showed that oligonucleotides containing these bases can be copied enzymatically³. However, as had been feared, the fidelity of copying was poor, perhaps because *iso-G* can exist in a tautomeric form that pairs with U. Now it has been shown that base pairs formed by κ with χ or with π do not suffer from this deficiency. Oligodeoxynucleotides containing κ and π were synthesized by chemical methods and shown to form helices almost as stable as those formed by oligomers containing only the standard base pairs. Next, it was confirmed that mispairing involving κ and π destabilizes these helices to about the same extent as mispairing involving the standard nucleotides.

DNA and RNA polymerases, it seems, are sensitive only to the external geometry of a base pair and indifferent to the arrangement of hydrogen bonds that maintains that geometry. Consequently, when the pyrimidine nucleotide analogue κ is incorporated into a template oligonucleotide, both T7 RNA polymerase and the Klenow fragment of DNA polymerase 1 will incorporate χ opposite it. The fidelity of replication using the DNA polymerase is good; no incorporation of dA or

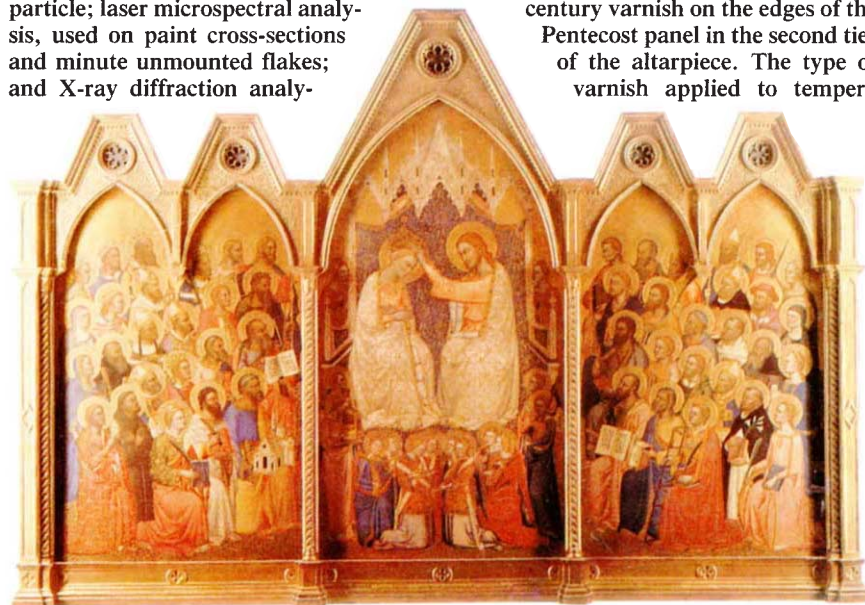
The science of art appreciation

WHAT analytical chemistry has to offer the art historian is one of the themes of *Italian Painting Before 1400*, an exhibition currently on show in London*.

The San Pier Maggiore Altarpiece (below), a highlight of the exhibition, is one of the largest and best-documented works produced in Florence in the second half of the fourteenth century. Research on it typifies the way in which scientific analysis of artists' materials moves from mere description to explanations of the painting technique and of artistic effect. Several methods are used for identifying pigments. Among them are energy-dispersive X-ray microanalysis carried out in the scanning electron microscope, where it is possible to analyse selectively a single pigment particle; laser microspectral analysis, used on paint cross-sections and minute unmounted flakes; and X-ray diffraction analy-

Many of the organic dyestuffs used to make artists' pigments deteriorate and fade so seriously that their presence and effect is not always easily discernible. In particular, the yellow lake, known as *arzica*, which is found in the San Pier Maggiore Altarpiece and other works in the exhibition, has been identified only rarely before. *Arzica* was notorious for its impermanence, even in the fourteenth century, yet artists continued to use it, especially for mixed greens, a hue in which the mediaeval palette was particularly weak.

It was through the use of another technique, gas chromatography – mass spectrometry, that one of the most significant finds of the exhibition emerged: the remnants of an original fourteenth-century varnish on the edges of the Pentecost panel in the second tier of the altarpiece. The type of varnish applied to tempera



sis, for fairly pure samples of crystalline pigments.

When applied to the altarpiece, the techniques revealed the range of the palette and confirmed that the painter Jacopo di Cione, fulfilled the terms of his contract by using the best materials available. The finest grade of the most expensive pigment, ultramarine blue (lapis lazuli), was explicitly specified and was reserved for the Virgin's robe — for mediaeval artists there was an established link between materials and meaning, intrinsic value and aesthetic appreciation.

Rather more interesting, however, is the identification of a range of red and yellow organic 'lake' pigments, which do not feature in the contract and were widely used on the altarpiece both in colour mixtures and as final glazes.

paintings and its visual effect is a subject much debated among both restorers and art historians. Tiny patches of a transparent orange substance were found directly overlying original paint, and these were identified as a sandarac-type resin dissolved in linseed oil, heat-bodied by boiling, with the addition of red lead to act as a drier. The components correspond with contemporary varnish recipes. As the catalogue accompanying the exhibition points out, our view of the original surface appearance should be reconsidered in the light of this discovery — the resin would produce a viscous and glossy varnish coating, which would also saturate the colours in a manner quite alien to current interpretations of how these paintings were intended to look.

Caroline Villers

* *Italian Painting Before 1400*, sponsored by Esso, is at the National Gallery, Trafalgar Square, London, until 28 February.

Caroline Villers is at the Courtauld Institute of Art, Somerset House, Strand, London WC2R 0RN, UK.

dG opposite κ was observed, and χ was not incorporated in positions opposite to bases other than κ if the full set of deoxynucleoside triphosphates was available. Transcription using T7 RNA polymerase is more error-prone, but this is not surprising because RNA polymerases are usually less accurate than DNA polymerases. The demonstration that the incorporation of χ in the template leads to the faithful incorporation of κ in the product would complete the proof that an extended genetic alphabet is possible.

The pyrimidine analogues discussed, with the exception of *iso-C*, are substances in which the base and sugar are joined by a carbon-carbon bond. *Iso-C* and *iso-G* do not form a completely satisfactory base pair (see above), which perhaps accounts for their 'rejection' by evolution. Perhaps the other base pairs were not used because methods for the prebiotic synthesis or early biosynthesis of the pyrimidine analogues were not available.

The incorporation of χ as one of the main constituents of a replicating nucleic acid might present special problems because, at near neutral pHs, the base carries a negative charge that is not present on the standard nucleosides. It would be interesting to study the stability of the

double helices formed by oligomers containing a substantial proportion of xanthosine. Incidentally, it is perhaps surprising that DNA and RNA polymerases accept χ TP as normal substrate, despite the extra charge that it carries.

Piccirilli *et al.* expect that the expansion of the genetic alphabet from four to six or more letters will enable them to incorporate functionalized monomers into oligonucleotides with extended catalytic capacity. This is an interesting suggestion, but it is not obvious to me that a similar result could not be achieved using chemically synthesized oligomers based on functionalized derivatives of the standard nucleotides. The authors also suggest that, in the long run, the genetic code can be extended artificially. How many amino acids can be coded for with triplets derived from a 12-letter alphabet? Don't forget to take account of wobble! $\square$

Leslie E. Orgel is at The Salk Institute for Biological Studies, PO Box 85800, San Diego, California 92138, USA.

1. Piccirilli, J.A., Krauch, T., Moroney, S.E. & Benner, S.A. *Nature* **343**, 33–37 (1990).
2. Rich, A. in *Horizons in Biochemistry* (eds Kasha, M. & Pullman, B.) 103 (Academic, New York, 1962).
3. Switzer, C.Y., Moroney, S.E. & Benner, S.A. *J. Am. chem. Soc.* **111**, 8322 (1989).

EVOLUTION

A case of male opportunism

Malte Andersson

THE idea that male secondary sex traits evolve through female choice has been controversial since Darwin¹ first proposed it. Over a century later it remains so, as discussed in a News and Views article four years ago². The two most debated explanations^{3–5} are Fisher's⁶ runaway process and various forms of indicator mechanism, both of which assume that male trait and female preference evolve together. On page 66 of this issue⁷, Ryan *et al.* report that neither applies to the case of the tungara frog (*Physalaemus pustulosus*). Females are attracted by a male call, but their preference has not evolved together with it; rather, the call seems to take advantage of a pre-existing frequency bias of the female ear. If similar mechanisms apply in many species, present views of sexual selection by female choice may have to be changed.

In some animals, female mate choice can be explained by advantages in the form of resources provided by the male, which enable the female to have more offspring⁸. In others, males contribute nothing but genes. For such animals, the fisherian explanation holds that the female preference evolves together with the preferred male trait in a process that becomes self-reinforcing. It can therefore bring the male trait to an extreme stage of

development. Indicator mechanisms, on the other hand, suggest that the female preference evolves together with a male trait that advertises high heritable viability.

Ryan *et al.*'s conclusion — that in the tungara frog the female preference remains unaffected by the evolution of the male call — was reached by comparing *P. pustulosus* with a closely related frog, *P. coloradum*. The two species show similar frequency tuning of the basilar papilla, the main sound receptor stimulated by the 'chuck' part of the male call. Yet *P. coloradum* does not produce chucks. The simplest explanation seems to be that the similar basilar papilla tuning in the two species has been inherited from a common ancestor. If so, the female preference for chucks in *P. pustulosus* has not evolved together with the male call. The call instead appears to be an opportunistic adaptation that takes advantage of the given bias of the female ear.

The results help explain why females in this species favour low-frequency calls, and why large males, which produce such calls, have a mating advantage. Audio-grams show that the sensitivity of the female basilar papilla is maximal for calls of lower frequency than those of the average male. Calculations based on properties of the female auditory system and the

frequency spectrum of male calls indicate that lowered fundamental frequency of the chuck call will stimulate females more effectively. For these reasons, sexual selection of the call characteristics and body size in male tungara frogs seems to stem, at least in part, from the properties of the female ear.

At least two questions are raised by these results. First, the fisherian model predicts that the female preference evolves along with the male trait in polygynous species — so why has this not occurred to any appreciable degree in the tungara frog? One possibility is that there has not been sufficient genetic variation in the preference. Second, what factors have selected for the present frequency tuning of the female (and perhaps male) ear? Further analysis of these aspects would be of interest.

Fisher formulated his runaway model of female choice to explain how sexual selection can lead to extreme sex traits that reduce male survival. The chuck call of the male tungara frog means the frogs risk predation by frog-eating bats⁹, and is therefore precisely the kind of sex trait that made Fisher (and Darwin) arrive at their ideas on sexual selection. But the evidence presented by Ryan *et al.* suggests that the chuck call has not evolved by runaway selection. Also, in some other animals, conspicuous male sex ornaments have been shown to be selected in other ways¹⁰. The complex process envisaged by Fisher will perhaps turn out to be less important than has been thought, and simpler mechanisms may account for even the most conspicuous secondary sex signals.

That is by no means certain, however. The possibility of genetic coevolution of the preference and the preferred trait has been supported by other studies, for example of ladybirds¹¹. Much more work remains to be done before a reliable general picture of the evolution of secondary sex traits by mate choice will emerge. $\square$

Malte Andersson is in the Department of Zoology, University of Gothenburg, PO Box 25059, 40031 Gothenburg, Sweden.

1. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (Murray, London, 1871).
2. Partridge, L. & Harvey, P. *Nature* **323**, 580–581 (1986).
3. Bradbury, J.W. & Andersson, M.B. *Sexual Selection: Testing the Alternatives* (Wiley, Chichester, 1987).
4. Kirkpatrick, M.A. *Rev. Ecol. Syst.* **16**, 43–70 (1987).
5. Pomiankowski, A.N. *Oxford Surv. Evol. Biol.* **5**, 136–184 (1988).
6. Fisher, R.A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
7. Ryan, M.J., Fox, J.H., Wilczynski, W. & Rand, A.S. *Nature* **343**, 66–67 (1990).
8. Alcock, J. *Animal Behavior: An Evolutionary Approach*, 4th edn (Sinauer, Sunderland, 1989).
9. Tuttle, M.D. & Ryan, M.J. *Science* **214**, 677–678 (1981).
10. Butcher, G.S. & Rohwer, S. *Curr. Ornithol.* **6**, 51–108 (1988).
11. Majerus, M.E.N., O'Donald, P., Kearns, P.E.W. & Ireland, H. *Nature* **321**, 164–167 (1986).

PLASMA PHYSICS

Plasmas everywhere

Richard D. Petrasso

ALTHOUGH described as the fourth state of matter, plasmas — gases of ionized atoms and molecules — are in fact the rule, not the exception, in the Universe. It is the clement conditions of Earth that make them seem a rarity. Exciting advances in space plasmas and in controlling laboratory plasmas, be they for controlled nuclear fusion or for X-ray lasers, were a feature of a recent wide-ranging meeting*.

Of the visible matter in the Universe, over 99 per cent consists of plasmas. Taking the case of the Solar System, the mass is totally dominated by a gravitationally confined sphere of plasma — the Sun. Although often considered to be electrically neutral mixtures of (ponderous) ions and (highly mobile) electrons, plasmas can also comprise a single charged species, such as trapped electrons, ions, positrons or antiprotons. Because plasmas, whether neutral or not, consist of such charged particles, they can support and react to electric and magnetic fields and exhibit a wide variety of collective modes, waves and instabilities. In fact these collective properties are exploited in beams of charged particles in particle accelerators and coherent radiation devices such as free electron lasers and gyrotrons.

Mirroring this diversity in character is the exceptional range of scales of plasmas (Fig. 1), from 10^{-3} to 10^6 K in temperature and from 10^2 m $^{-3}$ to 10^{32} m $^{-3}$ in density. At the high-density end of the spectrum, laboratory carbon/deuterium plasmas compressed by laser pressure (S. Nakai, Osaka Univ.) recently achieved a record density (around 600 g cm $^{-3}$). For the deuterium component, this is only a factor of three less than the proton number density (10^{32} m $^{-3}$) in the solar core. At the opposite end of the density spectrum, the

Voyager spacecraft observed tenuous (10^2 – 10^3 m $^{-3}$) high-energy protons in the magnetospheres of Saturn and Neptune (S. M. Krimigis, Johns Hopkins Univ.). For both planets, despite their intrinsic differences, the magnetospheric bulk proton temperature is 640×10^6 K (Fig. 2). Even at this elevated temperature, there exists a population of protons in the high-energy tail whose effective temperature significantly exceeds the bulk value. The fascinating challenge presented by these observations is to identify a heating or acceleration mechanism by which magnetospheric protons can reach such elevated temperatures and energies. Not only must a satisfactory theory account for this heating, but it cannot depend sensitively upon the angle between the planetary rotation and magnetic axis (about 1° for Saturn, and an unexpected 47° for Neptune). An equally fundamental problem is the origin of the magnetic dynamo that generates not only the solar but planetary magnetic fields — and hence the magnetospheres — for Saturn, Neptune, Uranus, Jupiter and Earth (see the News and Views article by H.K. Moffatt, *Nature* **341**, 285–286; 1989).

The most celebrated application of magnetically confined neutral plasmas is the attempt to generate controlled nuclear fusion through so-called 'tokamaks', toroidal plasma-containment devices in which a large magnetic field (2–12 tesla) and plasma current (50 kiloamps to 6 megamps) thread the torus thereby inhibiting the loss of hot plasma particles. One of the notable results was the near achievement at the Joint European Torus of plasma conditions required for equivalent fusion breakeven (M. Keilhacker, JET, Oxford). Breakeven is a conceptual landmark for fusion physics wherein the

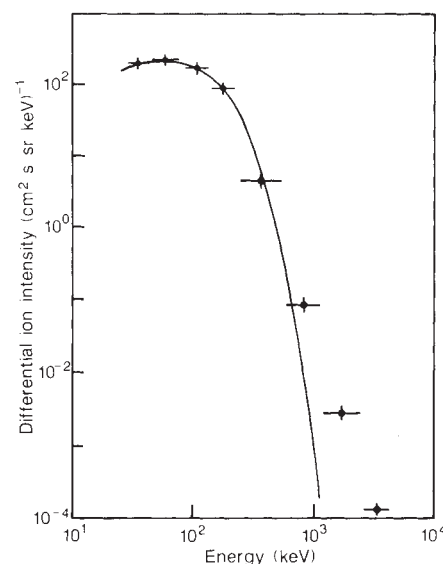


FIG. 2 Proton energy distribution for the magnetosphere of Saturn. The Maxwellian curve (solid line) is a fit to $T = 640 \times 10^6$ K (55 keV). Even for this remarkably high temperature, significant deviations still occur for proton energies above 600 keV. What mechanism generates this elevated temperature and high-energy, non-Maxwellian component? The proton energy distribution for Neptune's magnetosphere is strikingly similar to that of Saturn. (Courtesy of S. Krimigis, Johns Hopkins Univ.)

power from fusion reactions, usually assumed to arise in a fuel mixture of deuterium (D) and tritium (T), equals the input power needed to heat the plasma.

The figure-of-merit that measures the approach to breakeven, and also to the more meaningful goal of plasma ignition, is the triple product, $Tn\tau$, where T is the central ion temperature, n the central ion plasma density, and τ the global energy confinement time (the duration for which the plasma heat can be effectively trapped). Near the temperature peak in a deuterium discharge (Fig. 3), JET reached a record $Tn\tau$ of 1×10^{20} K m $^{-3}$ s for 0.1 s ($T = 250 \times 10^6$ K, $n = 3.7 \times 10^{19}$ m $^{-3}$ and $\tau \approx 1.1$ s). With a plasma fuelled by a deuterium–tritium mixture in a discharge similar to this one, JET would have momentarily obtained 80 per cent of breakeven. This would correspond to 13 MW of fusion power being generated at the peak temperature by the reaction $D + T \rightarrow n$ (14.1 MeV) + ^4He (3.5 MeV). For plasma ignition, a further improvement by a factor of 6 is needed for $Tn\tau$. Contrast this, however, with the 25-order-of-magnitude excess over the threshold value for D–T ignition found in the Sun ($T = 15 \times 10^6$ K, $n = 1 \times 10^{32}$ m $^{-3}$ and $\tau = 3 \times 10^{14}$ s).

JET's recent success largely stems from inhibiting carbon, oxygen, nickel and other contaminants from entering the plasma either at contact points between the plasma edge and internal tokamak

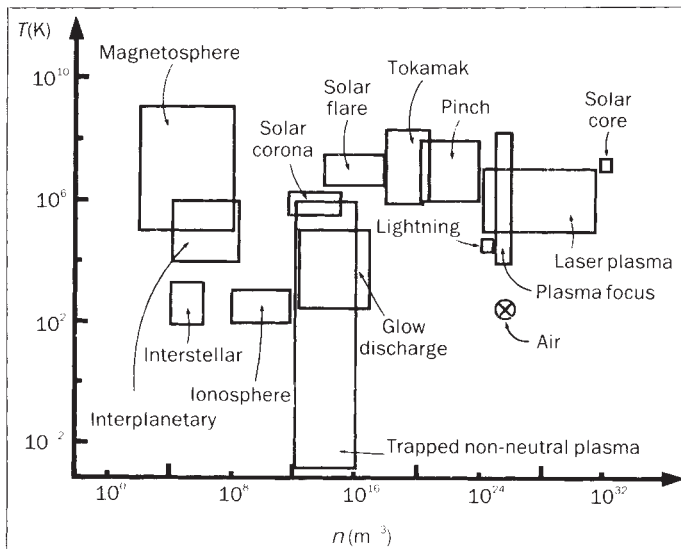


FIG. 1 A qualitative density-temperature map of plasmas. In density n the range spans 30 orders of magnitude, in temperature T , 12 orders. For reference, the $\otimes$ marks the density and temperature of air.

*American Physical Society, Division of Plasma Physics, annual meeting, Los Angeles, 13–17 November 1989.

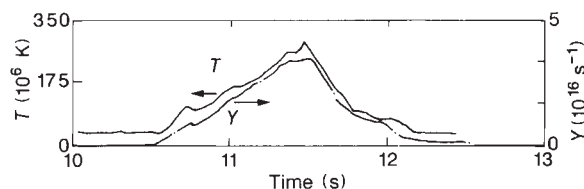


FIG. 3 Time evolution of central ion temperature (left axis, T) and D-D neutron rate (right axis, Y) for JET plasma discharge 20981 ($D-D \rightarrow n + {}^3\text{He}$). At 11.4 s the temperature peaks at 250×10^6 K, which is 17 times hotter than the Sun's core (15×10^6 K). Most importantly, the $Tn\tau$ product (see text) also peaks at this time. At 11.5 s, the temperature and $Tn\tau$ drop precipitously owing to an influx of carbon impurities. The principal aim for JET now is to eliminate such influxes. (Courtesy of M. Keilhacker.)

structures or from sputtering (ion-surface collision) processes. When too many impurities inadvertently enter the plasma, n and T dramatically decrease owing to fuel dilution and increased radiative losses. JET reduced the impurity levels by replacing carbon tiles in the tokamak with beryllium ones and coating the remaining interior surfaces with beryllium. In fact the sharp reduction in peak temperature (250×10^6 K) at 11.5 s (Fig. 3) is a direct result of the rapid influx of carbon from graphite tiles yet to be replaced.

In the domain of laser/plasma physics, two novel developments were, first, the initial demonstration of an X-ray laser operating at $44.83 \text{ \AA}$ (B. J. MacGowan, Lawrence Livermore National Lab.), which is just to the long wavelength side of the carbon K-shell X-ray absorption edge; and, second, the application of table-top high-power lasers (H. Milchberg, Univ. Maryland) to the generation of plasmas with densities equalling those of solids (10^{29} m^{-3}). The X-ray laser work uses two arms of the Nova laser at Lawrence Livermore to ionize a thin strip of tantalum (73 electrons) down to the closed-shell, nickel-like state (28 electrons). Many of these nickel-like ions are collisionally excited from the $3d$ to the $4d$ shell. The lasing action, about 400 ps in duration, occurs between $4d$ and $4p$ states. There is interest in this particular wavelength because such radiation could be applied to the holography of (carbon-rich) biological specimens. At wavelengths slightly longer than the carbon K-shell edge ($43.7 \text{ \AA}$) there is sufficient scattering in the specimen to achieve good contrast while still minimizing the adverse effects of specimen heating through absorption (R. A. London, M. D. Rosen & J. Trebes *Appl. Optics* **28**, 3397–3404; 1989). Experiments in the near future will concentrate on doubling the X-ray path length in the Ta plasma (currently sufficient to give laser amplification by e^4) by the addition of a normal-incidence X-ray mirror at one end of the lasing medium.

The application of short-pulse (sub-picosecond), table-top, millijoule lasers to irradiate solid targets is also an active area of research. This is partly a result of

being able to generate warm plasmas (10^3 – 10^6 K) at solid-like densities. When these lasers operate with a very small prepulse, the density scale length at the laser-plasma interface can be of the order of or less than the collisionless skin depth, $c/\omega_p \approx 200 \text{ \AA}$ (c is the speed of light, ω_p the plasma frequency). An intriguing feature of such plasmas is that they can probe regimes for which the plasma coupling parameter (Γ), the ratio of interparticle potential to kinetic energy, is of the order of 1 (for example, at $T \approx 10^5$ K and $n \approx$

10^{29} m^{-3}). Much of classical (not quantum degenerate) plasma physics concerns weakly coupled regimes for which $\Gamma \ll 1$. This is the case for (Voyager) space plasmas ($\Gamma \sim 10^{-13}$), tokamak plasmas ($\Gamma \sim 10^{-6}$), the X-ray laser plasmas ($\Gamma \sim 10^{-3}$) and even the solar core ($\Gamma \approx 0.1$). At the other extreme, recently achieved non-neutral trapped-ion plasmas have yielded $\Gamma \geq 10^4$ and are therefore strongly coupled. Thus plasmas with $\Gamma \sim 1$, such as those made by short-pulse lasers, form a pivotal bridge linking weakly and strongly coupled plasmas. $\square$

Richard D. Petrasso is at the Plasma Fusion Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

meter (Γ), the ratio of interparticle potential to kinetic energy, is of the order of 1 (for example, at $T \approx 10^5$ K and $n \approx$

CONSERVATION

What the little birds tell us

Jeremy J. D. Greenwood

THE consequences of the destruction of tropical forests are not always immediately apparent to citizens of the developed temperate zone. But now, in identifying a general decline in populations of birds that breed in eastern North America and winter in the neotropics, Robbins and colleagues¹ report strong evidence that deforestation in the tropics has a direct effect on the natural history of a temperate region. Two-thirds of the pairs of breeding birds in many North American forests are neotropical migrants² and American birdwatchers have been asking "Where have all the birds gone?". Once again, as in the controversy over pesti-

work, such as the North American Breeding Bird Survey (BBS)³. During the first 15 years of the survey (1965–79), no general decline in the populations of neotropical migrants was apparent, so Hutto⁴ has argued that reports of such declines result from local changes in the breeding areas — fragmentation of American and Canadian forests, for example. He has pointed out that many of the migrants are common in disturbed, non-forest habitats in winter, and so are unlikely to have been adversely affected by the destruction of neotropical forests.

Many studies do indeed show that forest fragmentation in the North American

Kevin Baker

breeding grounds particularly affects long-distance migrants and species living in the interiors of forests: their population densities are lower in small areas of woodland than in large areas, and at least some of the local declines have occurred only in such small fragments⁵. Why should long-distance migrants be particularly sensitive to forest fragmentation? Two clues are that a decline in a small woodlot which coincided with re-

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Bay-breasted warbler, *Dendroica castanea*, a typical long-distance migrant that breeds in northern forests and winters in the neotropics. Its breeding numbers declined by almost 16 per cent each year between 1978 and 1987.

cides in the 1950s and 1960s, birds are proving useful indicators of the widespread ecological effects of human activities. There has, however, been disagreement about the extent of the population declines and whether their causes are to be found on the summering or the wintering grounds.

We know a great deal about bird populations because thousands of volunteers are prepared to collaborate in survey

removal of local forests was reversed when the forests became re-established⁶, and that long-distance migrants have higher rates of return to breeding sites in subsequent years than do residents and short-distance migrants⁵. The second observation may simply reflect the generally higher survival rates of migrants than of residents; but perhaps the long-distance migrants, which do not return to their breeding area until shortly before nesting

must begin, have little time to find small and isolated parcels of woodland.

Terborgh² and others have argued that the many reports of local population declines are evidence of a widespread phenomenon, one caused by neotropical deforestation rather than by local effects in North America. Terborgh points out that the timing of the declines coincides with the large-scale clearance of neotropical forests (especially at higher altitudes, where migrants concentrate) rather than with northern deforestation. And he maintains that the species that have suffered most severely are the forest specialists.

Robbins and his colleagues¹ have now analysed the more recent BBS data. They show that, although populations of neotropical migrants throughout eastern North America were stable or increasing before 1978, there has been a general decline since then; this has not been paralleled in residents or short-distance migrants. And they confirm that it is the species that winter in forests which have been chiefly affected: species that winter in scrub remain common. Although habitat changes in the breeding areas are clearly important locally, we must now accept that neotropical deforestation is causing severe losses in the breeding birds of North America.

What of European parallels? The African-European bird migration system is very different to that in the Americas: Africa has less forest and is host to fewer migrants, not many of which winter in forests³. Furthermore, large reductions in the numbers of forest-breeding species probably occurred hundreds or thousands of years ago as the European forests were cleared for farming⁴. There have been recent declines in the populations of some migrant species, but these declines tend to coincide with changing rainfall patterns in Africa rather than with deforestation; indeed, some of them have been reversed as rainfall has improved⁵.

The difference between Old World and New World migrants warns us against making generalizations that are too broad and too simple. The declining neotropical migrants are mostly warblers but include an array of other species — orioles, buntings, tanagers, vireos, thrushes, flycatchers and cuckoos — and the precise reasons for the declines are probably not the same in all of them. Indeed, Robbins and his colleagues point out that, although neotropical forest destruction is undoubtedly a major general factor, fragmentation of forests in North America may also be important, at least for some species. A

critical test of whether tropical or northern forest destruction is the more important would involve analysis of the population trends of species that move from forest to non-forest habitats (or vice versa) when they migrate from temperate to neotropical regions. Unfortunately, there are too few of them. What we now need are intensive studies of the ecology of a range of species to check the details of the picture that the broad surveys have revealed. □

Jeremy J.D. Greenwood is at the British Trust for Ornithology, Beech Grove, Tring, Hertfordshire HP23 5NR, UK.

GEOCHEMISTRY

Back to the Earth's beginnings

H. Palme

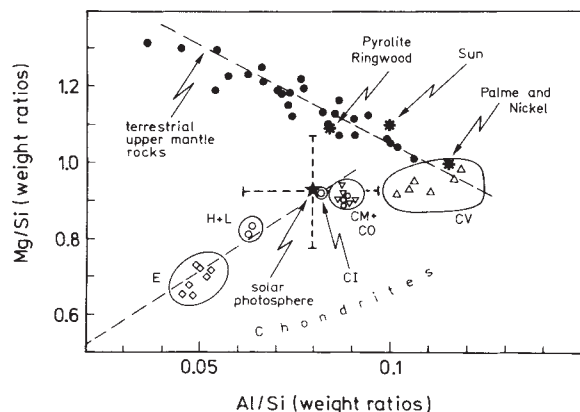
MODELS of the thermal and chemical evolution of the Earth are continually hampered by a lack of knowledge of the starting materials. The composition of minerals in the Earth's mantle must depend on both the mechanisms of fractionation that have occurred over geological time and the composition of the solar nebula from which the Earth was formed. A. E. Ringwood, in a challenging contribution¹, abandons common assumptions concerning the relationship between the terrestrial mantle, meteoritic materials and the composition of the Sun, to arrive at new ideas about the composition of the nebula from which the Earth accreted.

Continuous improvement in solar photospheric spectroscopy has provided us with a nearly complete abundance table for the chemical composition of the Sun. One group of meteorites, the type 1 carbonaceous chondrites (CI chondrites), with its most prominent member Orgueil, match the solar abundances for a surprisingly large number of elements² to within 10 per cent. These primitive meteorites are widely considered as being the best available sample for the average composition of the Solar System, failing only for very volatile elements, such as hydrogen, carbon, nitrogen, oxygen or the rare gases, and for the light elements lithium, beryllium and boron.

Consequently, CI chondrites have been used as a standard for the bulk composition of the Earth. However, the ratio of magnesium to silicon, two of the most abundant rock-forming elements, is significantly higher

(by about 20 per cent) in the upper mantle of the Earth than in the Orgueil meteorite, both ratios being marginally compatible with the solar ratio considering its larger error bars (see figure). Various attempts have been made to explain the obvious lack of Si in the upper mantle: loss of the more volatile Si during accretion of the Earth; partitioning of Si into the Earth's core; or a chemically fractionated terrestrial mantle with a Si-rich lower mantle that would balance the low Si content in the upper mantle.

It is the basic assumption — that the Mg/Si ratio of the bulk Earth equals that of the CI chondrites — that Ringwood



Variations in Mg/Si and Al/Si ratios of upper-mantle rocks (solid symbols; spinel ilherzolite xenoliths in alkali basalts) reflect the removal of variable amounts of Si- and Al-rich melts from the upper mantle. The composition of the primitive upper mantle (not depleted in Al and Si) must lie somewhere at the lower right end of the line defined by the upper-mantle rocks. Although there is some scatter in the Mg/Si ratios assumed for the upper mantle, all suggested compositions are higher than that of the CI-chondrites (meteoritic compositions, open symbols). One type of C3 (CV) carbonaceous chondrites would compositionally match an Al-rich upper mantle. The variability of the Mg/Si ratio in chondritic (unfractionated) meteorites is not understood. Solar ratios from Anders and Grevesse². Error bar for solar ratio was calculated from individual errors, applying gaussian error propagation. The figure is adapted from Jagoutz *et al.*³. Data for upper-mantle estimates: Ringwood¹; Sun²; Palme and Nickel⁴.

1. Robbins, C.S., Sauer, J.R., Greenberg, R.S. & Droege, S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7658–7662 (1989).
2. Terborgh, J.W. in *Migrant Birds in the Neotropics: Ecology, Behaviour, Distribution, and Conservation* (eds Keast, A. & Morton, E.S.) 21–30 (Smithsonian Institution, Washington DC, 1980).
3. Robbins, C.S., Bystrak, D. & Geissler, P.H. *The Breeding Bird Survey: Its First Fifteen Years, 1965–1979* (US Department of the Interior, Washington DC, 1986).
4. Hutto, R.L. *Am. Birds* **42**, 375–379 (1988).
5. Leck, C.F., Murray, B.G. & Swinebrod, J. *Biol. Conserv.* **46**, 145–157 (1988).
6. Askins, R.A. & Philbrick, M.J. *Wilson Bull.* **99**, 7–21 (1987).
7. Greenberg, R. in *Migrant Birds in the Neotropics: Ecology, Behavior, Distribution, and Conservation* (eds Keast, A. & Morton, E.S.) 493–504 (Smithsonian Institution, Washington DC, 1980).
8. Karr, J.D. in *Migrant Birds in the Neotropics: Ecology, Behavior, Distribution, and Conservation* (eds Keast, A. & Morton, E.S.) 529–543 (Smithsonian Institution, Washington DC, 1980).
9. Moreau, R.E. *The Palearctic-African Bird Migration Systems* (Academic, London, 1972).
10. Marchant, J.H., Hudson, R., Carter, S.P. & Whittington, P. *Population Trends in British Breeding Birds* (British Trust for Ornithology, Tring, Herts, 1990).

now challenges. Ringwood maintains that the upper-mantle Mg/Si ratio is representative of the bulk Earth. His arguments in favour of a chemically homogeneous terrestrial mantle are based on results of experiments on the partitioning of major and trace elements between high-pressure phases that are stable in the deep interior of the Earth and silicate melt.

Crystallization and gravitational settling of such phases from a completely or partially molten terrestrial mantle would preferentially incorporate certain elements, such as aluminium or hafnium. The residual mantle should then be deficient in these elements. Because samples from the upper mantle, accessible to laboratory analyses, do not show this depletion, Ringwood concludes that differentiation of the Earth's mantle has not occurred. Ringwood then postulates, contrary to assumptions that he and others have previously made, that the bulk-Earth (upper mantle) Mg/Si ratio, and not the CI ratio, is identical to the solar ratio.

Ringwood offers in support of his hypothesis systematic studies of the Mg/Si ratio in various objects in the Solar System. He concludes that the inner planets — Earth, Moon, Venus, Mars and the parent planet of a group of closely related basaltic and ultramafic meteorites (possibly derived from the asteroid Vesta) — could all have higher Mg/Si ratios than the Orgueil meteorite. On the other hand, low Mg/Si ratios are found not only in CI chondrites: there is evidence that interplanetary dust particles (IDPs) recovered from the stratosphere by U2 aircraft have similar or even lower ratios. Also, chemical data on the composition of dust particles from comet Halley obtained by European and Soviet spacecraft indicate a low Mg/Si ratio there. Because the inner planets including the Earth have the solar Mg/Si ratio according to Ringwood's hypothesis, CI-chondrites, IDPs and comets, having lower Mg/Si ratios, should be enriched in Si.

A common feature of CI chondrites, IDPs and comet Halley is that they are rich in volatile elements and it is therefore assumed that they formed at low temperatures in cooler outer parts of the solar nebula. Because Si is, under most conditions, more volatile than Mg, a high content would be consistent with high contents of volatiles.

Ringwood supposes that during formation of the terrestrial planets, volatile elements — including a small fraction of Si — would be lost and later recondense on material in the outer part of the solar nebula. The amount of Si lost would have to be so small that the terrestrial planets would still retain the solar Mg/Si ratio.

Although Mg/Si ratios can be reliably determined in CI meteorites, it may seem impossible to determine such a ratio precisely for a bulk planet from the analysis of

a few differentiated surface rocks. However, one must realize, as Ringwood points out, that small differences in this ratio may reflect large variations in the ratio of the two important Mg and Si bearing minerals, olivine, Mg_2SiO_4 , with an Mg/Si ratio of 2 and orthopyroxene, MgSiO_3 , with a ratio of 1. Basalts, partial melts extruded on the surface of a planet, are accessible to direct or remote chemical analysis. As the chemical composition of basalts is primarily determined by the mineralogy of their source region it is a sensitive indicator for the Mg/Si ratios in the deep interior.

As with all such models, some of Ringwood's assumptions are open to challenge. Most importantly, the opinions on the question of a chemically homogeneous terrestrial mantle are divided. D. Anderson (California Institute of Technology) has consistently pointed out that the velocity of seismic waves travelling through the mantle requires a more Si-rich lower mantle. Other modellers still believe they can accommodate melting and crystallization of the mantle with constraints obtained by the partitioning data of Kato *et al.*³ Furthermore, the possibility that some Si was removed from the mantle during the formation of the core cannot be excluded (see Laura Garwin's News and Views report⁴ for a discussion of this).

In addition, one may argue that a volatility-related 20 per cent excess of Si in CI chondrites, should result in correspondingly greater excesses of more volatile elements than Si, such as sulphur, sodium and potassium. The agreement between abundances in the solar photosphere and in Orgueil for these elements would then have no significance.

Although details of the model may be questioned, some kind of chemical stratification of the solar nebula is probably required to understand variations in the chemistry of meteorites, comets and the inner planets. The importance of Ringwood's approach lies in the attempt to relate the composition of the Earth to that of other Solar System materials and to try to understand chemical variations in terms of a uniform model of Solar System evolution. This type of modelling will become increasingly important in the future as new compositional data on Solar System objects become available. □

H. Palme is in the Max-Planck-Institut für Chemie, Saarstrasse 23, 6500 Mainz, FRG.

The cold smoulder

MOST big cities are unattractive places these days. As social harmony and civic pride decline, their inhabitants increasingly disfigure them with litter and graffiti, which depress civic pride even more. Daedalus now proposes a technical fix.

He has been inspired by those catalytic surfaces which burn up organic deposits. Many metallic and metal-oxide surfaces are excellent oxidation catalysts at elevated temperatures — they are used in those ovens which rapidly burn off grime and spilt food. Some carefully optimized surfaces can oxidize selected organics at 100 °C or even lower. So he is devising a catalytic pavement, the DREADCO 'Catwalk', to oxidize paper, plastics and so on, at room temperature.

As a low-temperature catalyst, Catwalk will absorb atmospheric oxygen very readily. The tricky problem is activating the organic litter lying on it. Daedalus points out that solid surfaces are always microscopically rough. They touch each other at only very few points, where the contact pressure is extremely high. The pressure-activated molecules at these points should therefore oxidize easily, by a form of catalytic stress corrosion. As fast as they burn away, the unsupported solid will subside upon their successors, which will oxidize in their turn.

Instead of the few milliseconds of contact time needed for high-temperature catalysis, this low-temperature variant will take hours or days to degrade a piece of litter. But ultimately anything organic — torn newspaper, discarded fast-food wrapper, cigarette end, painted slogan or offensive sticker — will slowly burn to ash, or partially oxidize to some rain-soluble mixture of aldehydes and polyols.

Once perfected, 'Catwalk' and its vertical counterpart 'Catawall' should make a city almost vandal proof. The graceless, resentful citizens will chuck their garbage on its streets, or degrade its hideous architecture still further with graffiti, only to see their exertions rapidly nullified. They may rejoice in this endless renewal of their creative canvas; but just possibly they may reform, and take pride in their new, smart environment.

Catwalk and Catawall will have intriguing side-effects. Shoe soles and tyres will rapidly oxidize away, and may need to be replaced by inorganic or silicone-based substitutes. Street-corner loungers will lose the shoulders of their jackets, and dogs will have to wear booties. Hygiene will be wonderfully enhanced. The universal dusty detritus of life, hair and fibre and food fragments and general organic dirt, and the bacteria and cockroaches and mice that live on it, will simply oxidize and disappear. The urban environment will be almost surgically clean.

David Jones

1. Ringwood, A.E. *Earth planet. Sci. Lett.* **95**, 1–7 (1989).
2. Anders, A & Grevesse, N. *Geochim. Cosmochim. Acta* **53**, 197–214 (1989).
3. Kato, T. *et al. Earth planet. Sci. Lett.* **89**, 123 (1988).
4. Garwin, L. *Nature* **338**, 19–20 (1989).
5. Jagoutz, E. *et al. Proc. 10th Lunar Planet Sci. Conf.* 2031–2050 (1979).
6. Ringwood, A.E. *Composition and Petrology of the Earth's Mantle* (McGraw Hill, New York, 1975).
7. Sun, S.-S. *Geochim. Cosmochim. Acta* **46**, 179 (1982).
8. Palme, H. & Nickel, K.G. *Geochim. Cosmochim. Acta* **49**, 2123–2132 (1985).

Life and the Sun's lifetime

SIR—Carter¹ has noted the remarkable coincidence between the timescale of biological evolution on Earth and the lifetime of the Sun. He has used this fact, in conjunction with the anthropic principle, to argue that extraterrestrial civilizations are exceedingly rare, even if conditions favourable to the development of life are common. Barrow and Tipler² have discussed this point further.

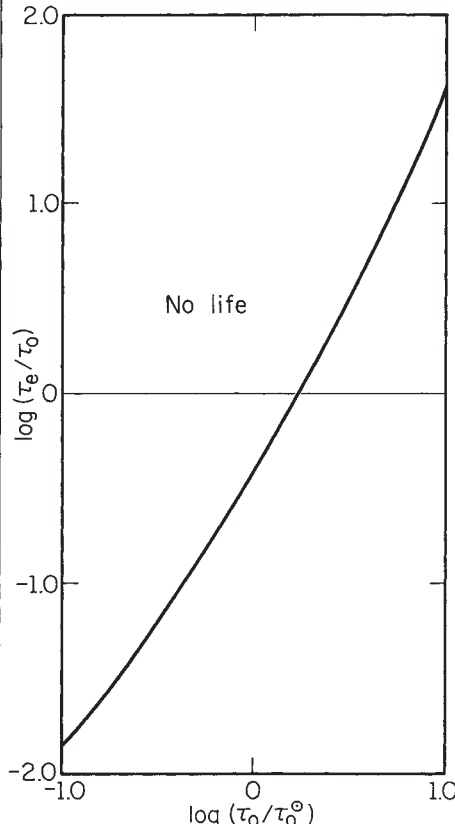
Carter's conclusions rely on the supposition that the two timescales are independent, but we argue that there does exist a definite relationship between the stellar mass (and therefore its lifetime) and the timescale for the development of life on a suitable planet. If such a relationship exists, a coincidence of the type $\tau_e \approx \tau_0$ can no longer be dismissed on the basis of *a priori* probabilities. Here τ_e is the timescale of biological evolution, taken for simplicity to correspond to the appearance of land life, and τ_0 is the lifetime of the star. We assume that the criterion for life to develop is that oxygen must reach some critical level in the atmosphere of a planet^{3,4}.

This simple conjecture is based on the idea that filtering out of ultraviolet radiation is necessary to make the planet inhabitable; it is certainly a reasonable assumption for carbon-based life. To obtain a quantitative relationship, we consider planets similar to the Earth in terms of location, radius and initial composition, that belong to planetary systems around different stars. This is sufficient for our purpose, because the statement of Carter and Barrow and Tipler that there is no physical connection between the timescales of stellar and biological evolutions means essentially that if we were to keep the Earth in place and change the central star (within reasonable limits), this would not affect biological evolution. This approach means that we consider only a limited fraction of parameter space, but we are interested only in general trends.

Two main phases in the development of oxygen in a planet's atmosphere can be identified. In the first, oxygen released from photolysis of water vapour removes all reduced gases (such as CH_4 and NH_3) from the atmosphere. In the second, the amounts of O_2 and O_3 in the atmosphere become large enough to reduce the intensity of ultraviolet radiation to tolerable levels. In simulations of the evolution of the Earth's atmosphere⁴, the first stage lasted about 2.4×10^9 yr and the second lasted about 1.6×10^9 yr. This simulation was tailored to fit the conditions on Earth, requiring that a stable atmosphere would develop over ~ 4.5 Gyr; we emphasize, however, that our conclusions do not depend at all on the details of these simulations. The timescale for the first phase is

inversely proportional to the intensity of radiation in the range 1,000–2,000 Å (refs 3–5). For the second phase, we assume that the ultraviolet radiation must be reduced by extinction in the atmosphere roughly to the present-day level on Earth: $I_{\text{uv}}^{\text{planet}} \approx I_{\text{uv}}^{\text{Earth}} \approx 10^{-4} I_0^\odot$, where $I_0^\odot$ is the incident intensity on the upper atmosphere. We consider stars with lifetimes in the range $0.1 \leq \tau_0/\tau_0^\odot \leq 10$.

In the calculations, we use (1) a mass-luminosity relation for main-sequence stars⁶ ($L/L_\odot \approx (M/M_\odot)^{3.45}$); (2) a mass-



The ratio of the timescale for life to develop, τ_e , to the stellar main-sequence lifetime, τ_0 , as a function of τ_0 . Life cannot develop for $\tau_e > \tau_0$.

radius relation⁶ ($R/R_\odot = (M/M_\odot)^\alpha$, with α in the range 0.6–1.0 for this mass range; and (3) the fraction of the total energy emitted in the range 1,000–2,000 Å (ref. 6). Expressing the ultraviolet extinction as $I_{\text{uv}}^{\text{planet}} = I_0 e^{-kx}$, where I_0 is the intensity impinging on the planet's atmosphere, k is the absorption coefficient and x is the depth of the atmosphere, we use relations (1)–(3), combined with the assumption about the necessary level of extinction, to calculate, for each stellar lifetime τ_0 , the timescales of the two phases in the evolution of oxygen in the planet's atmosphere and therefore the timescale for the development of life.

The results are presented in the figure. Clearly there exists a relationship between τ_0 and τ_e , τ_e being a monotonic

function of τ_0 . Although the quantitative details are uncertain, as many complicating factors have not been considered, we thus consider the case for a qualitative relationship to be established. Therefore, the possibility that $\tau_e \approx \tau_0$ cannot be excluded on the basis of *a priori* probabilities.

Carter's argument was based on the absence of any such relation, in which case all intervals (of equal length) along a vertical line passing through $\log(\tau_e/\tau_0^\odot) = 0$ in the figure are equally probable. Once a τ_e – τ_0 relation is established, however, this is no longer the case. Why $\tau_e(\tau_0)$ passes through the particular point $(0.4 \tau_0^\odot, \tau_0^\odot)$, corresponding to the case of the Earth, is a question that depends on the detailed physics and biology of the development of life. But if we accept the fact that $\log(\tau_e/\tau_0)$ is a monotonically increasing function of τ_0 (even if different from the one in the figure), then because the number of stars increases with increasing τ_0 (for a Salpeter initial mass function⁷), the probability of finding $\tau_e/\tau_0 \approx 1$ is the highest.

M. LIVIO*†
A. KOPELMAN*

*Department of Physics,
Technion, Haifa 32000,
Israel, and

†Department of Astronomy,
University of Illinois,
Urbana, Illinois 61801, USA

1. Carter, B. *Phil. Trans. R. Soc.*, **A310**, 347–363 (1983).
2. Barrow, J.D. & Tipler, F.J. in *The Anthropic Cosmological Principle* (Clarendon, Oxford, 1986).
3. Berkner, L.V. & Marshall, L.C. *J. Atmos. Sci.*, **22**, 225 (1965).
4. Hart, M.H. *Icarus*, **33**, 23–39 (1978).
5. Brinkman, R.T. *J. geophys. Res.*, **74**, 5355–5368 (1969).
6. Allen, C.W. *Astrophysical Quantities* (Athlone, London, 1973).
7. Salpeter, E.E. *Astrophys. J.*, **121**, 161–167 (1955).

Mummy RNA lasts longer

SIR—Svante Pääbo suggests^{1–3} that nucleic-acid fractions isolated from mummified archaeological remains are composed of depolymerized and chemically modified DNA, with a minor proportion (less than about 1 per cent) of unmodified double stranded DNA.

The presence of modified DNA is inferred from two main kinds of evidence. First, HPLC elution profiles of the hydrolysis products of ancient nucleic acids show extremely reduced thymine peaks compared with modern standards and second ancient polynucleotides are highly alkali-sensitive.

We have recently analysed the composition of the nucleic-acid fractions isolated both from 1,000- and 3,300-year-old seeds of maize and cress from Peru and the Thebes necropolis^{4,5} and from a tissue fragment isolated from an unidentified Egyptian mummy (courtesy of Lanfredo Castelletti, Museo Civico 'Paolo Giovio',

Como). The techniques we used were anti-DNA antibody binding, HPLC fractionation, ribonuclease digestion and molecular hybridization by reverse-phase blotting. We found that IgG antibodies, purified from the sera of patients with systemic lupus erythematosus and specific for single and double-stranded DNA, precipitated only up to 1 per cent of the total nucleic acids. HPLC analysis, on a C18 reverse-phase column, of the formic acid hydrolysis products of the different samples, revealed very reduced thymine peaks but sharp uracil peaks were detected immediately after cytosine peaks. Moreover, uracil appears to be present as unidentified peaks in Pääbo's chromatograms^{1,3}. Our results also show that brief treatments with DNase-free RNase degrades most of the polynucleotides. Finally, in the case of the maize seeds, radiolabelled ancient nucleic acids hybridize strongly with cloned RNA genes, whereas only very faint signals are produced with unexpressed, highly repeated sequences such as maize H2a.

Taken together, these data suggest that the nucleic acids in ancient samples are

composed mainly of RNA. The reason, we suspect, is as follows: for a short time span after the death of the organism, nucleic acids are degraded by endogenous nucleases. Sooner or later, however, all enzymatic activities cease, leaving only chemical and physical environmental factors to determine the further degradation of both DNA and RNA. Under these conditions we have no reason for believing that RNA is degraded much faster than DNA and because ribosomal RNA is by far the most abundant nucleic acid in the cell, it may still be present in significant amounts long after all the DNA has vanished.

FRANCO M. VENANZI

FRANCO ROLLO

*Dipartimento di Biologia Cellulare,
via F. Camerini 2,
I-62032 Camerino,
Italy*

1. Pääbo, S. *J. Arch. Sci.* **12**, 411–417 (1985).
2. Pääbo, S. *Cold Spring Harb. Symp. quant. Biol.* **51**, 441–446 (1986).
3. Pääbo, S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1939–1943 (1989).
4. Rollo, F., Amici, A., Salvi, R. & Garbuglia, A.R. *Nature* **335**, 774 (1988).
5. Rollo, F. *Theor. appl. Genet.* **71**, 330–333 (1985).

No more neutrino cold dark matter

SIR—Measurements of the width of the Z boson were reported recently from the LEP and SLC colliders^{1–5}. These data place limits on the number of light neutrino species. We would like to draw attention to one of the implications of the results, namely that they rule out both Dirac and Majorana neutrinos as viable candidates for cold dark matter, a possible form of the dark matter that is known to exist in galactic halos (see, for example, refs 6–11).

Precise masses have not been obtained for any of the three known neutrino species, but experiments require¹² $m_{\nu_e} < 18$ eV, $m_{\nu_\mu} < 250$ keV and $m_{\nu_\tau} < 35$ MeV. Moreover, limits on the mean cosmological density will be exceeded by relic neutrinos with mass 100 eV $< m_\nu < 1$ GeV (refs 6–11), implying that if one of these three known species of neutrino constitutes the dark matter, it must be the 'hot' variety. But simulations of large-scale structure favour 'cold' dark matter as the most consistent means of accommodating present data on galaxies. To supply such cold dark matter typically requires particles in the GeV mass range (or of the axion type). A supersymmetric particle is probably the theoretically preferred GeV dark-matter particle, but a fourth species of neutrino of either the Dirac type (four spin states) or Majorana type (only two spin states, because Majorana particles are their own antiparticles) was the first candidate considered for cold dark matter and continues to draw much attention^{6–11}.

The four LEP experiments have now

collected in excess of eleven thousand Z bosons^{1–5} and, by measuring the cross-section near the peak of the Z resonance, have been able to constrain the number, N , of light ($m_\nu < m_Z/2$) neutrino species. The results are^{1–5} OPAL: $N = 3.12 \pm 0.42$; ALEPH: $N = 3.27 \pm 0.30$; L3: $N = 3.42 \pm 0.48$; DELPHI: $N = 2.4 \pm 0.64$. The MARK II detector at SLC (ref. 2) gives $N = 2.8 \pm 0.6$. Taking a weighted average over all these experiments, one finds the number of neutrino species $N = 3.13 \pm 0.20$.

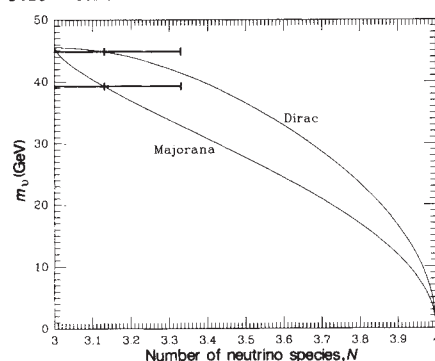


FIG. 1 Mass of hypothetical fourth-generation neutrino as a function of the measured 'number' of neutrino species, assuming the deviation from $N=4$ is due to the mass suppression in the Z width. The data points are the combined results from SLC and LEP plotted separately for the Dirac and Majorana cases.

The width Γ of Z decay into neutrinos is a function of the mass m_ν of the neutrino. For neutrinos with masses greater than $m_Z/2$ the decay rate, r , is zero, whereas for lighter neutrinos we have, for

the Dirac case,

$$r_D = \frac{\Gamma(Z \rightarrow \nu_D \bar{\nu}_D)}{\Gamma(Z \rightarrow \nu \bar{\nu})_{\text{massless}}} = \left(1 - \frac{4m_\nu^2}{m_Z^2}\right)^{1/2} \left(1 - \frac{m_\nu^2}{m_Z^2}\right) \quad (1)$$

and for the Majorana case,

$$r_M = \frac{\Gamma(Z \rightarrow \nu_M \nu_M)}{\Gamma(Z \rightarrow \nu \bar{\nu})_{\text{massless}}} = \left(1 - \frac{4m_\nu^2}{m_Z^2}\right)^{3/2} \quad (2)$$

Setting $N=3+r_D$ or $N=3+r_M$, the LEP/SLC measurement can be taken as a measurement of the mass of a new (fourth-generation) neutrino. As it stands, the LEP/SLC measurements are consistent with three neutrino species (that is, $m_\nu > m_Z/2$, if it exists at all), and any deviation from three may, of course, be due to physics other than a new neutrino species. Nevertheless, for the present purpose one can attribute any deviation from $N=3$ to a fourth species, and can then solve equations (1) and (2) for its mass m_ν as a function of the measured 'number' of neutrino species. The result is shown in Fig. 1 for both the Dirac and Majorana cases. The LEP/SLC data, shown as the points with error bars in Fig. 1, imply that masses of a fourth species above ~ 40 GeV for Dirac neutrinos and above ~ 30 GeV for Majorana neutrinos are consistent with the data, and that masses far below these values are unlikely.

The mass of a neutrino species determines its relic abundance in the Universe today^{6–11}. Figure 2 shows the relic density as a function of N for both Dirac and Majorana neutrinos. The variable h is the Hubble constant in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and Ω is the present average mass density in neutrinos divided by the critical density. The approximate, but accurate, method used in calculating the relic abundance is described in refs 13 and 14. Astrophysical observations constrain the dark matter to satisfy $\Omega_{\text{DM}} \geq 0.1$, and the age of the Universe constrains $\Omega_{\text{DM}} h^2 < 1$. The allowed region is delimited by dashed lines in Fig. 2.

It is clear from the LEP/SLC data that a relic abundance consistent with galactic and cluster dark matter is unlikely in both the Dirac and the Majorana cases. For the Dirac case, however, it is possible for the relic abundance to be higher than shown in Fig. 2, as there may be more neutrinos than antineutrinos and the relic abundance calculation applies only to the case of no cosmic asymmetry. However, two groups^{15,16} have reported results from modified double-beta-decay experiments which rule out Dirac neutrinos as the primary constituent of the dark matter in precisely this mass range (> 12 GeV). One must conclude that it is unlikely that either Dirac or Majorana neutrinos in the GeV

mass range make up the bulk of the dark matter and that the long tradition of considering neutrino cold dark matter should come to an end. Constraints can also be imposed from the LEP experiments on supersymmetric dark-matter candidates¹⁷.

Finally we should mention some possible loopholes in our line of reasoning. The errors in the LEP and SLC measurements are still large, and future

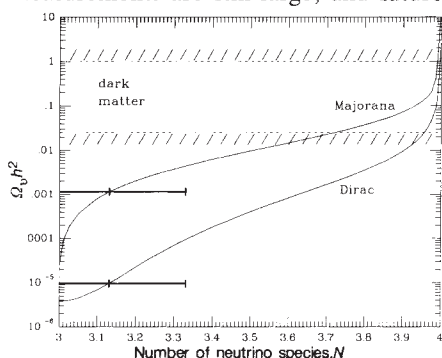


FIG. 2 Relic abundance of hypothetical fourth-generation neutrino as a function of the measured 'number' of neutrino species for the Dirac and Majorana cases. The data points are the combined LEP and SLC results. The region where the neutrino relic abundance is appropriate to supply the bulk of the dark matter ($\Omega_\nu = 0.1$ and $h = 0.5$ to $\Omega_\nu = 1$ and $h = 1$) is shown by dashed lines.

results could move the data point enough to make it consistent with the relic abundance limits. Note from Fig. 1, however, that to allow the canonical dark-matter neutrino mass of 5–10 GeV, N must be close to 4, a value several standard deviations away from the present measurement. Also note that, contrary to claims in the literature, the relic abundance of extremely massive neutrinos ($m_\nu \gg \text{TeV}$) must again approach the critical density, if the neutrinos remain elementary particles¹⁸. The existence of such particles is, of course, not constrained by the LEP experiment and they may make up the dark matter. Perturbation theory breaks down long before this limiting mass is reached, however, so it is not clear that a

theory with such heavy particles makes sense. Note that additional entropy generated in the early Universe could change the results of our relict abundance calculation, but in this case the abundance would be smaller than we claim, and neutrino cold dark matter would be even more unlikely.

KIM GRIEST
JOSEPH SILK

Center for Particle Astrophysics and
Department of Physics and Astronomy,
University of California,
Berkeley, California 94720, USA

Shedding light on PCR contamination

SIR—The most pernicious problem plaguing the widely used technique of polymerase chain reaction (PCR) is contamination of reagents with previously

Effect of ultraviolet irradiation on PCR. A standard protocol⁶ was used to prepare 20- μl PCR mixes, containing 50 mM KCl, 10 mM Tris-HCl pH 8.3, 1.5 mM MgCl_2 , 0.1% (w/v) gelatin, 200 μM each dNTP and 0.1 μM of each of two PCR primers that amplify a 750-base pair region of human factor IX complementary DNA. Linearized factor IX plasmid (6-kilobases in length) was added so that the amount of target DNA was as indicated below. The samples were irradiated for 5 (lanes 2–8) or 20 min, (lanes 10–16) and then denatured for 10 min. *Taq* polymerase (0.5 U) was added, and 30 cycles of PCR performed⁶. Lanes: S — standards, *Hind*III digestion of ϕX174 ; 1, 9 — no DNA and no ultraviolet; 2, 10 — no DNA; 3, 11 — 3 pg DNA; 4, 12 — 30 pg target DNA; 5, 13 — 300 pg target DNA; 6, 14 — 3,000 pg target DNA; 7, 15 — 30,000 pg DNA; 8, 16 — ultraviolet irradiation before addition of 3 pg non-irradiated target DNA (positive control). Method of irradiation. The 20- μl samples in a clear 0.5-ml polypropylene microcentrifuge tube were placed in direct contact with a glass platform of the Fotodyne 1000 transilluminator. Irradiation was performed with 254- and 300-nm ultraviolet bulbs in the standard instrument configuration.

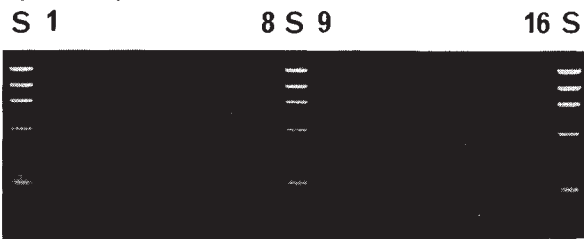
amplified material^{1–3}. Care is essential in the routine handling of amplified samples, and heroic efforts are required if PCR from single cells is to be safely performed^{3,4}. The following experiment illustrates how treatment with ultraviolet light directly before adding the template DNA can eliminate most sources of contamination.

Between 3 and 30,000 pg of target DNA was added to 20 μl PCR mix, which was then irradiated for 5 or 20 minutes with a combination of 254- and 300-nm bulbs and subsequently amplified with 30 cycles of PCR. As shown in the figure, even when the quantity of DNA was 10⁵-fold (5 min) or 10⁶-fold (20 min) more than the minimum quantity of DNA that we routinely detect with PCR⁵, no amplified product was seen. DNA added after irradiation of the PCR mix could still be amplified, indicating that irradiation of the PCR mix does not compromise the ability of the reagents to mediate efficient amplification. Irradiation with the 300-nm transilluminator alone, though considerably less

effective than if the 254-nm bulb is also used, should still be adequate for the levels of contamination that routinely occur in practice.

In additional experiments, concentrated stock solutions of buffer ($\times 10$), oligonucleotide primers ($\times 10$), deoxytriphosphates ($\times 7$), or *Taq* polymerase ($\times 40$) were contaminated with 1.5 $\mu\text{g } \mu\text{l}^{-1}$ of target DNA and irradiated for 5 minutes with 254- and 300-nm ultraviolet light. In each case, the contaminant could not be amplified whereas 0.015 $\mu\text{g } \mu\text{l}^{-1}$ of target DNA added after irradiation was efficiently amplified.

It was initially surprising that the 5-min ultraviolet treatment reduced the efficiency of amplification from double-stranded DNA by more than 100,000-fold but preserved the functional integrity of the single-stranded oligonucleotide primers in the PCR mix. However, this seems generally to be the case because 5-min irradiation of four additional unrelated



pairs of oligonucleotide primers did not reduce the intensity of the product amplified from 250 ng of human genomic DNA.

GOBINDA SARKAR
STEVE S. SOMMER

Department of Biochemistry
and Molecular Biology,
Mayo Clinic/Foundation,
Rochester, Minnesota 55905, USA

1. Kwock, S. *Amplifications* **2**, 4 (1989).
2. Williams, J.F. *BioTechniques* **7**, 762–766 (1989).
3. Gibbs, R.A. & Chamberlain, J.S. *Genes Dev.* **3**, 1095–1098 (1989).
4. Li, H. *et al. Nature* **335**, 414–417 (1988).
5. Sommer, S.S., Cassady, J., Sobell, J.L. & Bottema, C.D.K. *Proc. Mayo Clinic* **64**, 1361–1372 (1989).
6. Sarkar, G. & Sommer, S.S. *Science* **244**, 331–334 (1989).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references.

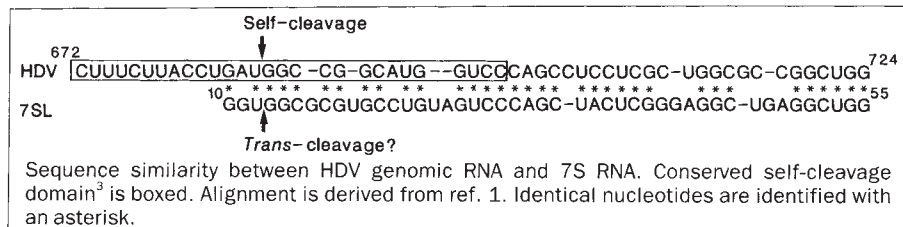
1. ALEPH collaboration: Decamp, D. *et al. Phys. Lett. B* **231**, 519–529 (1989).
2. OPAL collaboration: Akrawy, M. Z. *et al. Phys. Lett. B* **231**, 530–538 (1989).
3. DELPHI collaboration: Aarnio, P. *et al. Phys. Lett. B* **231**, 539–547 (1989).
4. L3 collaboration: Adeva, B. *et al. Phys. Lett. B* **231**, 509–518 (1989).
5. SLC results: Abrams, G. S. *et al. Phys. Rev. Lett.* **63**, 2173–2176 (1989).
6. Vysotskii, M. I., Dolgov, A. D. & Zeldovich, Ya. B. *JETP. Lett.* **26**, 188–190 (1977).
7. Lee, B. W. & Weinberg, S. *Phys. Rev. Lett.* **39**, 165–168 (1977).
8. Kolb, E. W. & Olive, K. A. *Phys. Rev. D* **33**, 1202–1203 (1986).
9. Scherrer, R. J. & Turner, M. S. *Phys. Rev. D* **33**, 1585–1589 (1986).
10. Srednicki, M. *et al. Nucl. Phys. B* **310**, 693–713 (1989).
11. Steigman, G. A. *Rev. Nucl. Part. Sci.* **29**, 313–338 (1979).
12. Aguilar-Benitez, M. *et al. Phys. Lett. B* **204**, 1–486 (1988).
13. Griest, K. & Seckel, D. *Nucl. Phys. B* **283**, 681–705 (1987); **B** **296**, 1034–1036 (1988).
14. Griest, K. & Sadoulet, B. in *Dark Matter in the Universe* (eds Galeotti, P. & Schramm, D. N.) (Kluwer, Dordrecht, The Netherlands, 1989).
15. Ahlen, S. P. *et al. Phys. Lett. B* **195**, 603–608 (1987).
16. Caldwell, D. O. *et al. Phys. Rev. Lett.* **61**, 510–513 (1988).
17. Barbieri, R. *et al. Phys. Lett. B* **195**, 500–506 (1987).
18. Griest, K. & Kamionkowski, M. Preprint FERMILAB-Pub-89/205-A (1989).

Delta virus as a cleaver

SIR—Negro *et al.*¹ recently reported sequence complementarity between hepatitis delta virus (HDV) RNA and human 7S RNA, the essential RNA component of the signal recognition particle (SRP), and suggested that HDV RNA exerts its pathogenic effect by base-pairing with the 7S RNA. We would like to suggest that instead of simply pairing to 7S RNA, HDV RNA might subsequently inactivate it by catalytic cleavage.

catalytic activity on substrates such as 7S, supplied in *trans*.

In this model, HDV cleavage would remove the 5'-terminal 12 nucleotides of 7S RNA, within the region of SRP required for translational arrest of secretory and membrane-associated proteins. Because HDV replicates in the nucleus^{1,6}, cleavage of 7S could occur before assembly of the mature ribonucleoprotein. RNA-catalysed cleavage of 7S RNA



In infected cells, both genomic and anti-genomic HDV RNA can be detected. *In vitro*, multimeric copies of both genomic and anti-genomic HDV can self-cleave²⁻⁴. A minimal contiguous self-cleavage domain (approximately 100–130 nucleotides) has been described for both strands and is apparently unrelated to the well characterized 'hammerhead' self-cleavage domains found in a variety of small RNAs⁵. The HDV catalytic domains of genomic and anti-genomic RNAs contain similar, approximately 25-nucleotide, sequences⁵.

Interestingly, a subset of the genomic self-cleavage domain, including the cleavage site, is found in 7S RNA (nucleotides 10–31 in the figure). This suggests that the 7S sequence could interact with the self-cleaving catalytic centre in HDV, resulting in cleavage of 7S at residue 12. Because the self-cleavage domains of the genome and anti-genome are similar in sequence, the same interaction is possible for the anti-genomic RNA. We suggest that both strands of HDV RNA, not just the anti-genomic strand, may play essential roles in the pathogenesis of HDV.

The capacity of catalytic RNAs to operate on substrates in *trans* has been recognized for several years^{6,7}. In group I intron-derived RNA enzymes (ribozymes), substrate selection occurs through base-pairing between the substrate and a sequence in the ribozyme's catalytic core⁷. We suggest a similar mechanism for HDV substrate recognition. The processed HDV genome, although no longer a substrate for self-cleavage, may retain

would prevent translational arrest of secretory and membrane proteins, causing ectopic expression of these proteins in the cytosol. If this is correct, HDV may be thought as an infectious ribozyme.

Sequence similarity between HDV and 7S RNA indicates additional intriguing possibilities. By binding components of the SRP, HDV may deplete the cell of SRP proteins or even facilitate its own export from cells by exploiting the host protein-transport machinery. Finally, as a large portion of the conserved HDV self-cleavage domain is present in the 7S sequence, 7S RNA may itself possess catalytic activity.

BENJAMIN YOUNG*

BRIAN HICKE*†

*Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, Boulder, Colorado 80309-0347, USA, and

†Medical Scientist Training Program University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

Bone-conducted sound

SIR—Kermack¹, in his recent News and Views article about hearing in early mammals, refers to the lightness of the middle-ear ossicles as a feature contributing to the ability of modern mammals to detect much higher frequencies than birds or reptiles. He does not, however, mention what I believe to be the main advantage conferred by the mammalian system of ossicles — the attenuation of bone-conducted sound. Studies by Ernst Bárány providing strong evidence for this function² were published in full in 1938 but have not received the attention they deserve: this function is not mentioned in

any account of auditory physiology (admittedly only in English) that I have read.

Bárány proposed that in a bird or reptile with only a single ossicle (the columella auris, connecting the eardrum directly to the oval window), bone-conducted sound is generated mainly as follows. When the skull undergoes vibration with a component along the columella, the inertia of the columella causes it to move relative to the skull, thus causing movements of the inner-ear fluids in the same way as if it were being displaced by sound vibrations impinging on the eardrum. In the mammalian ossicle system, however, the head of the malleus acts as a counterweight, bringing the centre of mass of the whole ossicular chain very close to the axis about which the incus and malleus are pivoted relative to the skull. Thus, when the skull vibrates, the whole ossicular chain moves with it — the footplate of the stapes does not move relative to the oval window, and the inner ear fluids are not displaced.

Bárány demonstrated this mechanism in humans by elegant measurements of bone-conducted sound, using a null method in which the auditory sensation was brought to a minimum by adjusting the amplitude and phase of airborne sound applied simultaneously. He showed that the enhancement of bone conduction when mass is applied to the eardrum (a phenomenon discovered by his father Robert Bárány³) is due to the centre of mass of the ossicles being shifted off the axis of the pivot.

Bárány makes the point that, in quiet surroundings, bone-conducted sound generated by breathing, chewing, creaking in joints and so on, forms the main background noise above which external sounds have to be detected, so that the range at which, for example, a prey animal can detect a predator will be increased by reducing this background. On the other hand, the function customarily attributed to the ossicles (that is, improving the match between airborne sound and the inner-ear fluids by the 3:2 reduction in the lever action of malleus and incus) would be equally well achieved by a 33 per cent reduction in area of the footplate of stapes (or columella), and would therefore provide no selective advantage in evolving the complex system of ossicles.

It would be interesting to know whether the arrangement of the quadrate and articular bones in *Morganucodon* and in other early mammals was such as to balance out, even if imperfectly, the generation of bone-conducted sound.

ANDREW HUXLEY

*The Master's Lodge,
Trinity College,
Cambridge CB2 1TQ, UK*

1. Negro, F., Gerin, J. L., Purcell, R. H. & Miller, R. H. *Nature* **341**, 111 (1989).
2. Sharmeen, L. M., Kuo, M. Y.-P., Dinter-Gottlieb, G. & Taylor, J. *J. Virol.* **62**, 2674–2679 (1988).
3. Kuo, M. Y.-P., Sharmeen, L., Dinter-Gottlieb, G. & Taylor, J. *J. Virol.* **62**, 4439–4444 (1988).
4. Wu, H. N. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 1831–1835 (1989).
5. Uhlenbeck, O. C. *Nature* **328**, 596–600 (1987).
6. Gowans, E. J. *et al. Virology* **167**, 274–278 (1988).
7. Zaug, A. J., Been, M. D. & Cech, T. R. *Nature* **324**, 429–433 (1986).

1. Kermack, K. *Nature* **341**, 568–569 (1989).
2. Bárány, E. *Acta oto-lar. Suppl.* **26** (1938).
3. Bárány, R. *Msch. Ohrenheilk. Lar-Rhinol.* **44**, 549 (1910).

A deeper shade of green

Stephen Cotgrove

The Global Environmental Movement. By John McCormick. Belhaven Press: 1989. Pp. 259. £22.50, \$35. Distributed in the United States by Indiana University Press.

In 1661, the naturalist John Evelyn was deploring the air pollution from coal burning which made the City of London resemble "the Court of Vulcan... or the Suburbs of Hell". But man-made environmental disaster was by no means a new phenomenon, even then. The abandonment of Sumerian cities nearly 3,700 years ago, overgrazing and deforestation resulting in the soil erosion of the hills of Attica, and the collapse of Mayan civilization through over-population, are but a few early examples.

By the second half of the nineteenth century, rising concern had spawned several protectionist groups. In Britain, the Society for the Protection of Animals was founded in 1824. The impact of industrialization on towns and countryside led to a growing criticism of its challenge to the moral and social order. The National Trust and the Council for the Protection of Rural England are but two monuments to growing public concern. In the United States, anxiety focused on the preservation of the wilderness and the great national parks, with the founding of the Sierra Club in 1892. By contrast, and to some extent in conflict with these ideals, the conservationist movement sought to promote the sustainable and 'rational' exploitation of natural resources.

The Global Environmental Movement documents the unfolding story of the gradual growth in environmental consciousness and concern until it reaches global proportions with the emergence of international agencies and cooperation after the Stockholm Conference of 1970-72. McCormick's book is scholarly and well researched, as is to be expected from a text which has its origins in a post-graduate dissertation. But McCormick's strength is also his weakness: he is long on description and short on analysis.

In Britain, three million people are now members of environment groups, "making the movement the biggest in British history" (p. viii). McCormick recognizes that this 'movement' is of great complexity. There is little in common between members of the National Trust and Greenpeace save a concern for some aspect of the environment. To describe all these groups as members of the same movement is to underplay the deep differences in goals and values, in perceptions of what constitute threats to the environment, and in what are considered to be the appropriate strategies of environmental protection.

Although McCormick acknowledges such complexities, the issues raised are barely addressed and certainly not treated in any depth. The absolutely crucial distinction is between 'deep' and 'shallow' ecology (or deep and pale greens). These two poles of the 'movement' differ dramatically both in their perceptions of the

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

United, but for how long? Environmental groups demonstrating against nuclear weapons.

threats facing the environment, and in their prescriptions for cure. To take one example — the exponential growth in human numbers compared with the arithmetical increase in food supply, first noted by William Petty and subsequently popularized by Thomas Malthus 150 years later. It was not until 1972 that the publication of two influential books again raised the spectre of disaster. To the threat of catastrophe from the population explosion, the *Blueprint for Survival* and the Club of Rome's *Limits to Growth* added dire warnings that exponential economic growth was unsustainable. Such views were heavily criticized and even ridiculed by the scientific and political establishment at the time. *The Doomsday*

Syndrome by John Maddox, for example, was a powerful and authoritative critique. Other critics were less temperate — epithets such as "eco-hysteria", "eco-maniacs" and "eco-nuts" were hurled at the "doomsters".

All of this is faithfully recorded by McCormick. What is a little disappointing is that he stops short of coming to grips with the problems and issues raised by such varied responses to environmental threats. Why are there such dramatic differences in perceptions of environmental dangers and in prescriptions for action? In his final chapter, McCormick begins to outline an explanation, but sketches it in only lightly and briefly. He notes the assertion that the roots of an environmental crisis are to be found in the dominant culture of industrial society — a coherent ideology which includes faith in science and technology, the subjection of nature and its exploitation as a resource, and, above all, a deep commitment to economic growth by government. Those (the great majority) who see the world from this perspective are markedly less likely to perceive environmental dangers, and markedly more likely to believe that market forces and science and technology will provide solutions.

McCormick is optimistic that a fundamental change in attitudes is under way. But no amount of concern over litter and waste disposal, driving fuel-efficient cars and conserving energy will address the challenge raised by the deep greens — the possible exhaustion of the world's stock of mineral resources in a tiny fraction of the duration of man's life on Earth, and the doubling of the world's population in the next 30 years. Above all, the pressures for continuing economic growth on which democracies depend for their stability, and the aspirations of the developing world, may well be inexorable.

The attitudes, beliefs and values which constitute the dominant culture of industrial societies are deeply rooted in upbringing, experience and interests. To see the world as the deep greens see it requires something akin to a religious conversion. It is such unproven and largely untestable assumptions that structure the way we 'see' the world and interpret the evidence. Above all, values are not testable in the laboratory. And what is reasonable and rational action to maximize one value — protecting nature — may seem irrational to those for whom economic values are self-evident benchmarks. It could be that it is the world view of the deep greens which is 'realistic'. But it may take a crisis to challenge the dominant industrial culture, with its bias against the environment. By then, it may be too late to prevent irreversible damage. □

Stephen Cotgrove is Emeritus Professor of Sociology at the University of Bath, Caverton Down, Bath BA2 7AY, UK.

Higher sight

Glyn Humphreys

Synesthesia: A Union of the Senses.

By Richard E. Cytowic. Springer-Verlag: 1989. Pp. 354. DM148, £53, \$75.

A CHARACTER in a popular British television programme experiences pink days, yellow hopes and dreams that are all the colours of the rainbow. In some people, such associations between abstract concepts and imageable colours seem to occur involuntarily and to be both vivid and memorable. The condition of experiencing such involuntary and perceptible cross-associations is called a synesthesia.

Syndromes such as synesthesia are rarely dealt with in psychological science. Many basic questions have simply not been put: are synesthetic experiences real or merely examples of expansive and metaphorical use of language; what is the cause of such experiences; what are the relations between different forms of synesthetic experience, and between these experiences and other elaborate associations (as when people develop prodigious and individual mnemonic strategies)?

Luria touched on some of these points in *The Mind of the Mnemonist* (Basic Books; London, 1968), but subsequently they have all too often been subsumed under the banner of romantic neurology. To some extent this has occurred because current theoretical accounts of brain and behaviour have little to say about 'super-abundant' states. Syndromes in which particular behavioural functions are impaired (such as reading, writing or recognition of visual objects) can be understood in terms of the lesioning of one or more components in the complex system underlying the behaviour. Syndromes reflecting behavioural patterns over and above those that normally occur are less easy to relate to the operation of identifiable processing components. This can be taken as an indictment of current theorizing. Rigorous study of such syndromes, in which basic questions are answered, could immeasurably enrich our understanding of the relationships between brain and behaviour.

In his book, Richard Cytowic attempts to provide the first complete account of synesthesia. But problems abound, not least the definition of the syndrome. Cytowic cites many examples of what he classes as synesthetic experience: in many cases, people experience the perception of colour or of simple visual shapes (lines, squares, circles or simple random patterns) in response to particular stimuli — music and numbers seem two common elicitors. In other instances, taste, smell or tactile experiences are reported. But what of artistic experience and expression? Cytowic includes a chapter on synesthesia

and art, incorporating an interview in which David Hockney discusses his use of colour in opera design. This would be fine, if only the lines between empirical finding and speculation were clearly demarcated. It is a pity that they are not, by clouding the boundaries, Cytowic fails to provide the groundwork for future studies.

He does, however, make many useful observations. For instance, synesthetic experiences such as cross-modal colour associations seem to take a consistent form, and are often associated with similar stimuli, in different people. They are experienced from an early age, involving simple rather than complex shapes or known objects. The preponderance of synesthetes are female, often from families where other members report similar experiences. In one case studied in some detail by Cytowic and his colleagues, the subject experienced a marked decrease in cortical blood flow during synthetic experiences which correlated closely to variations in the synesthetic experience produced by the intake of ethanol and amyl nitrate. Such observations suggest that there is a real perceptual phenomenon underlying the experience. But without attempts to draw boundaries between different types of synesthetic experiences, or to establish empirical means of separating synesthesia from metaphorical expression, the sceptical reader is left unconvinced. I longed to know what would happen to some of the synesthetes in standardized tasks of experimental psychology, such as colour-matching of numbers (would they be impaired if numbers are drawn in colours other than those with which they are normally associated?) or the Stroop colour word-naming task. Such evidence is sadly lacking.

Cytowic proposes that synesthetic experiences arise from activity in limbic structures in the brain under conditions where cortical activity is decreased. This is an interesting idea, and consistent with the cerebral blood-flow work. But no sooner does Cytowic discuss "the hippocampus model for synesthesia" (p. 174) than he proposes that "the neural real estate for synesthesia resides in the left (cerebral) hemisphere". It is very hard at times to hold on to exactly what is being proposed. This is indicative of the over-expansive writing style that hides rather than illuminates the syndrome.

Many perceptual phenomena are treated in a cavalier way. Early on, Cytowic raises the example of colour constancy to argue that the world is not as we see it. Well yes, but then colour constancy is not an hallucination; rather it is based on the constraining influence of neighbouring surfaces in the visual field, as the author later makes clearer when he discusses Land's retinex theory of colour vision. In his discussion of 'blindsight', Cytowic argues that even when patients cannot

detect the presence of a shape in their blind visual field, they can nevertheless show good shape discrimination. In fact, Weiskrantz in his book *Blindsight* (Oxford University Press, 1986) showed that such patients have good orientation rather than shape discrimination. The list is extensive. Yet hidden beneath the florid writing are the glimmerings of an interesting phenomenon. Syndromes such as synesthesia ought to be the subject of more rigorous scientific study. If Cytowic prompts us towards that end, then his book will have accomplished no mean feat. □

Glyn Humphreys is in the Cognitive Science Research Centre at the University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

Cruel world

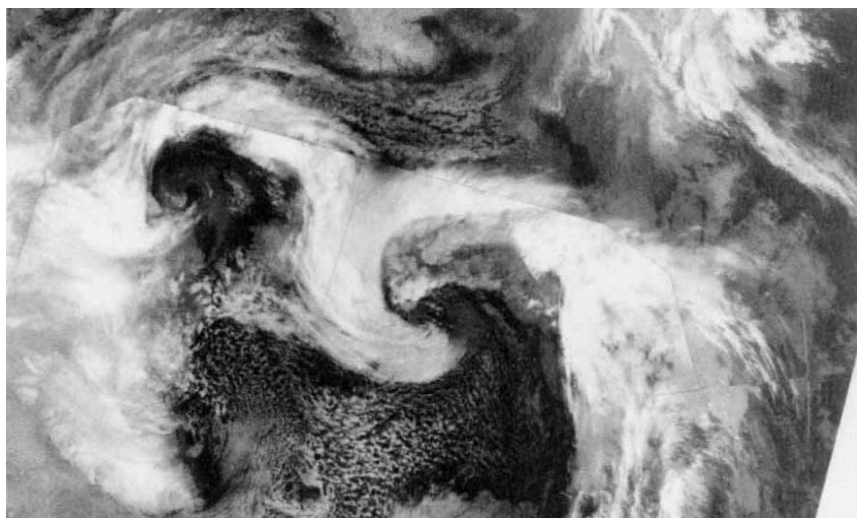
Andrew Holmes-Siedle

My Life with the Printed Circuit. By Paul Eisler. Associated University Presses, 440 Forsgate Drive, Cranbury, NJ 08512, USA/25 Sicilian Avenue, London WC1A 2QH. 1989. Pp.170. £13.95, \$29.50.

PAUL Eisler invented the printed circuit, a revolutionary method of manufacturing electronic hardware, in 1936. His adventures make instructive reading, especially for the engineer and innovator. Out of a detailed, clearly written story of a technical life (the useful bibliography is by Mari Williams) emerge many contrasts between the outlook of an engineer and of a businessman. This is not a hard-luck story, but it certainly is a cautionary tale.

Eisler's is not a plaintive account, for he has been widely honoured for his inventions and he has made a fair living as a research manager and consultant. He does not seem particularly bitter that many others have made millions by flouting his patents, but in this book he goes further than most to describe in detail the sheer sleaziness of the business practices that he encountered. I do not get the feeling that this nastiness is particular to the electronic industry — it is probably transferred unmodified from other manufacturing industries. The reader may be angered by the contrast between the grubby practices described and the clean technical promise of Eisler's invention (and of electronic systems as a whole). It is sad to find the same old grubbiness in that shining new world.

Few people realize that the printed circuit was invented long before the transistor. The invention was conceived in 1936 when Eisler, an Austrian engineer forbidden to take employment in Britain, was working on inventions in his small room in Hampstead, north London. A printing firm took up the printed circuit idea because the technique could be



Active front — stormy skies over the North Atlantic are shown in this composite satellite picture, from *Spacious Skies* by R. Scorer and A. Verkaik. Published by David and Charles, £20, \$35.

expected to lead to compactness and a saving of manufacturing labour in the wiring of electronics, including the consumer radios then on the market.

Then the war came and compactness was at a premium. Eisler's idea was taken up by the United States for anti-aircraft shell fuses, which turned out to be very effective against German and Japanese aircraft. Very strong patents, written by Eisler in 1943, were granted after the war. Unfortunately, almost simultaneously with the granting, the US National Bureau of Standards issued a brochure and organized a major conference on printed circuits. This must have confused many potential users and certainly gave those who wanted it the excuse to say "Oh, I thought this technology was owned by the US Government!". In fact, Eisler had assigned very strong rights to an English company, Technograph. In his method, by far the most practical for making printed circuits, a copper film is formed on a surface and then photolithographically etched into a pattern.

With the transistor came an opportunity to make electronic circuits very small indeed. It was the combination of the printed circuit and the transistor which enabled radios to be made pocket-sized, light and surprisingly tolerant of being dropped. The stakes in the ownership of the printed circuit increased greatly.

The lessons of Eisler's book revolve around the form that such high-powered industrial competitions take and the place in such a struggle of the legal rights to a key invention. Sadly, one lesson is that the rights of a case do not always matter. Litigation is so expensive that often only the powerful can uphold a right. Eisler's account gives detail and colour to this lesson and illustrates one facet particularly well. Innovators know that, with a new idea, they have to be somewhat generous with free help at the beginning. They do not expect anyone to buy a

licence — or even consulting time — until they have acquired a feel for what the invention will offer. In return, the originator hopes to spread the knowledge of his device in the right circles and, if possible, receive recognition of his key role and of the knowhow he has contributed.

Eisler's employers, astute businessmen themselves, did indeed give free help in several cases. Eisler comments in one case: "... expertise was made available by us to enable a large company to exploit ... a vast growth area ... the big organisations obtained the developments made by a small, weak firm free of charge. Factors of fairness or even recognition did not apply." In another case (telephone circuits), the printed circuit was seen by the manufacturers as a threat to their own, more futuristic proposals. As a result, funding by the post office was effectively stifled. In a third market, (specialized aircraft components), Eisler's employers were threatened with punitive action if they tried to work their way in. Proof that Eisler's rights were genuine is given by the fact that his employers gained a reasonable income from licences with many US companies, including Western Electric, IBM, General Motors and Kodak.

The charm of Eisler's story lies partly in the unvarnished, unemotional writing and the truly bizarre experiences mixed with the hard work. Here is a man who did things of which he is rightly proud and who is now using his powers of description and analysis to tell the story. Eisler concludes by describing the period when he exploited other inventions, and by looking back and asking whether the life of the inventor is worth the struggle. His book is a raw but substantial piece of industrial history — a portrait of an unpleasant manufacturing industry that rings true to me. □

Andrew Holmes-Siedle is consultant to Fulmer Research Ltd, Stoke Poges, Berkshire SL2 4QD, UK.

Light patches

Mary Osborn

An Atlas of Antigens: Fluorescence Microscope Localization Patterns in Cells and Tissues. By C. D. Ockleford. *Stockton Press: 1989. Pp.323. £85, \$125.*

THIS atlas is a serious effort to provide an easily accessible reference collection of immunofluorescence micrographs. It has strong and weak points. Its strength lies in an overview of antigen arrangement in cells and in a few selected tissues, in a good introductory methods chapter and in the references that it provides to the original literature.

The arrangement of about 200 antigens is illustrated in three colour plates and numerous black and white micrographs. Antigens are classified rather arbitrarily as nuclear, cytoskeletal, cell-surface, cytoplasmic organellar, diffuse cytoplasmic, extracellular matrix or as pathological targets. They are then, rather surprisingly, further subdivided within these sections alphabetically rather than by the structures they decorate. Thus the novice may not recognize when antibodies against two different proteins decorate the same structure, especially as the atlas switches between different cells in culture, and between specialized cell types and tissue sections from diverse species, seemingly at random. The result sometimes is like finding maps of London mixed up in sections on Asia, America and Australia.

A student with no experience in interpreting such micrographs may at times become hopelessly lost, as immunofluorescence reveals only the antigen-harboured structure recognized by the antibody against a dark background. Thus more line drawings or phase micrographs to illustrate exactly what structures or cell types are decorated, and more text to show which antigens occur in multigene families, would have been helpful. The quality of the micrographs is sometimes good (actin) and sometimes terrible (keratin, vimentin, extracellular-matrix proteins). In some instances a more stringent selection of micrographs (or of the laboratories supplying them) would have been useful. The pathology section is the weakest part of the book.

The atlas is, at least currently, neither the archival repository of information nor the easy reference source that its author hoped for. But construction of such an atlas is no easy task, and the many different and sometimes rather beautiful patterns present in the pages of this one may intrigue the newcomer to cell biology. □

Mary Osborn is at the Max-Planck Institute for Biophysical Chemistry, 3400 Goettingen, FRG.

Interior décor

Gregory Houseman

Mantle Convection: Plate Tectonics and Global Dynamics. Edited by W. R. Peltier. Gordon and Breach: 1989. Pp. 881. \$198, £116.47.

MANTLE convection has become one of the most exciting fields in the Earth sciences because it is at the heart of the basic physical and chemical configuration of our planet. It is a field that has seen great progress and intense disagreement in the past 20 years. An enormous volume of literature has been produced — pity help the newcomer who alone tries to find a way through the maze of papers expressing conflicting, and indeed contradictory, views. This volume makes a valuable contribution to the task of clearing a way through the minefield.

The editor has assembled substantial review papers by specialists in each of the fields of convection theory, seismic imaging, high-pressure physics, global geochemistry, mantle viscosity, numerical models of convection, geoid analysis and planetary accretion. The result is not necessarily a balanced overview of mantle convection, but it is relatively complete in that most of the lines of evidence that have been brought to bear on the question of mantle convection are covered. Some interpretations are more controversial than others; the obvious disadvantage of each topic being reported by only one or two groups of authors is that the interpretations presented don't necessarily reflect consensus with other groups working in that field. Although parts of the volume can be criticized on these grounds, I think that the net result is very successful. This volume could well be used as the starting point for a stimulating series of seminars on the subject of mantle convection and Earth structure.

An outsider might expect that, in 881 pages of well-documented scholarly review papers contributed by leading figures, there would at least be some consensus on the basic geometrical configuration of the convecting layer. Does the mantle consist essentially of a single-convecting layer, or of two layers separated by a density difference with a (probably leaky) interface about 700 kilometres deep? Paradoxically, one of the most basic parameters of the subject is one of the most difficult to determine. This question runs as a theme through most of the book, and although the balance of opinion seems to come down in favour of whole-mantle convection, some of the authors seem to find, as I do, that the evidence may have more than one interpretation. My guess is that the best prospect of an unambiguous resolution to this question lies with direct measurement

of the physical properties of silicates at the high pressures representative of lower mantle conditions.

One could turn to many of the problems discussed in this volume to illustrate the difficulties involved in applying an imperfect dataset to obtain perhaps poorly resolved parameters. The measurement of mantle viscosity is one of the most basic measurements relevant to mantle convection. Since Haskell's original observations in 1935, there have been numerous analyses as the database has grown. Lambeck, in his book *Geophysical Geodesy* (Oxford University Press, 1988), describes a vertical viscosity stratification significantly different to that summarized

here by Peltier. Is this difference perhaps a measurement of lateral viscosity variation in the mantle? The vertical viscosity structure inferred from analysis of the long-wavelength geoid is different again, requiring a very large increase in viscosity from the upper mantle to the lower.

I highly recommend this book as a survey of the present state of play in mantle convection, although the field has some way to go before an unequivocal description of the Earth's interior can be presented. □

Gregory Houseman is in the Department of Earth Sciences, Monash University, Clayton, Victoria 3168, Australia.

Detecting decay

C.K. Brain

Recent Vertebrate Carcasses and their Paleobiological Implications. By Johannes Weigelt. Translated by Judith Schaefer. University of Chicago Press: 1989. Pp. 188. Hbk \$68.95, £47.95; pbk \$22.95, £15.95.

THE justification for an English translation of a book that was published in German more than 60 years ago is the recent upsurge of interest in the discipline that has become known as taphonomy. Taphonomy deals with the rather macabre topics of death, decay, burial and fossilization, and its adherents are the palaeo-detectives of science who interpret the most meagre of clues in their reconstruction of disastrous events in the lives of long-dead organisms. One of the first and most distinguished of these palaeo-detectives was Johannes Weigelt (1890–1948), professor of geology and palaeontology at Martin Luther University, Halle-Wittenberg and founder of the Museum for Earth Science of Central Germany at Halle. His liberally illustrated book, written in the rambling style of a European naturalist, is widely respected as a seminal work in taphonomy but, until now, has not been available in English. Judith Schaefer's translation does full justice to the original text.

Taphonomy, coined by the Russian palaeontologist I. A. Efremov in 1940, is somewhat more inclusive than "biostratigraphy", the term Weigelt proposes in his book. The discipline, which has been remarkably successful in allowing the reconstruction of circumstances of life and death of early hominids and other populations, was formally constituted at a symposium held in 1976 at Burg Wartenstein, Austria, under the auspices of the Wenner-Gren Foundation for Anthropological Research. The resulting volume from that meeting, *Fossils in the Making*

(University of Chicago Press, 1980), was co-edited by Anna K. Behrensmeyer, the mother of taphonomy, who also contributes the foreword to the present volume, together with Catherine Badgley.

Weigelt's book is essentially a report written on his return to Germany after 16 months of intensive fieldwork on the Gulf Coast of the United States, where his mission had been to document processes of vertebrate death and burial to determine their relevance to fossil preservation. In Texas, Weigelt focused his attention on Smithers Lake, a body of water in Fort Bend County, whose level was prone to fluctuate a good deal, intermittently covering and exposing a vast assemblage of carcasses along its southern shore. The year, 1924, was appropriate for Weigelt's purposes: severe drought and icy winds killed more than a million head of cattle near Smithers Lake, together with fish, turtles and alligators. Weigelt scrutinized these disintegrating carcasses and found features already familiar to him among fossils in European museums. He concluded that many spectacular "death assemblages" from the remote past had resulted from analogous catastrophic events.

In his conclusion, Weigelt wrote that his aim had been to "stimulate interest in a method . . . that our sciences of geology and paleontology can continue to use to great advantage, and in so doing, win new friends". He would surely be gratified to find how many new friends he now has. □

C.K. Brain is honorary Professor of Zoology at the University of the Witwatersrand, and is Director of the Transvaal Museum, PO Box 413, Pretoria 0001, South Africa.

New in paperback

- *Understanding Human Evolution* 2nd edn by F. E. Poirier. Published by Prentice Hall, price £27.60, \$35.60.
- *Western Civilization in Biological Perspective* by S. Boyden. Published by Oxford University Press, price £15, to be published in the United States January 1990. See review in *Nature* 331, 670 (1988).

Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet

Joseph A. Piccirilli, Tilman Krauch, Simon E. Moroney & Steven A. Benner

Laboratory for Organic Chemistry, ETH Zurich, CH-8092 Switzerland

A new Watson–Crick base pair, with a hydrogen bonding pattern different from that in the A·T and G·C base pairs, is incorporated into duplex DNA and RNA by DNA and RNA polymerases and expands the genetic alphabet from 4 to 6 letters. This expansion could lead to RNAs with greater diversity in functional groups and greater catalytic potential.

LIFE on Earth seems to have passed through several episodes, including one where ribonucleic acid (RNA) was the sole genetically encoded component of biological catalysis^{1–5}. Evidence that RNA can act as a catalyst^{6–10} supports this view and now influences research in areas as diverse as prebiotic chemistry¹¹ and virology¹². Furthermore, models presuming the existence of an RNA world can now rationalize many of the metabolic details of contemporary organisms¹³, details which otherwise appear arbitrarily idiosyncratic. One important conclusion from these models is that RNA catalysed many types of reactions before proteins emerged¹⁴.

Even those who find this picture convincing recognize that proteins assumed at some point most of the catalytic roles performed earlier by RNA in the RNA world. This implies that proteins are selectively advantageous as catalysts, and the grounds for this advantage deserve exploration. One difference between the two type of macromolecules is the greater number of building blocks, and accompanying chemical functionalities, available to natural proteins compared with natural RNA molecules. Conversely, the limited functionality present on natural oligonucleotides constrains the chemist attempting to design catalytically active RNA molecules, in particular, RNA molecules that catalyse the template-directed polymerization of RNA.

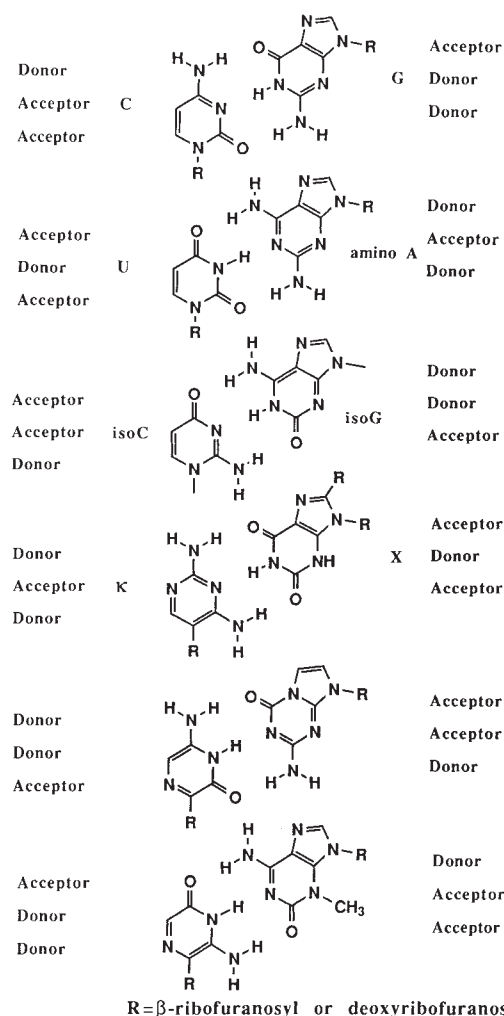
Additional base pairs could relax these constraints, especially if their hydrogen-bonding patterns differed from those in the A·T and G·C base pairs, as novel hydrogen-bonding schemes would allow additional base pairs to be incorporated enzymatically at specific positions in an RNA molecule^{14,15}. If functionalized, such additional bases might also allow the incorporation by transcription of functional groups directly into RNA; the remaining unfunctionalized building blocks could then control secondary structure without introducing over-functionalization and attendant nonspecific catalysis. Although this strategy for increasing the catalytic versatility of RNA has apparently never been pursued by nature, it has special appeal to chemists, who often wish to reproduce the properties of natural systems without precisely duplicating natural chemical structures.

The first problems in implementing this strategy are technological: the design, synthesis and study of new bases that can be incorporated by DNA and RNA polymerases uniquely opposite complementary bases in a template. This technology is needed as RNA molecules are, at present, best obtained by transcription of synthetic DNA using RNA polymerases¹⁶. Furthermore, it would be desirable if DNA polymerase could also

incorporate any newly designed bases selectively in a growing DNA chain, as this would permit the amplification of oligonucleotides containing additional nucleoside 'letters'. We report here a new base pair that meets both of these requirements.

New base pairs

The geometry of the Watson–Crick base pair can accommodate at least six mutually exclusive hydrogen-bonding schemes (Fig. 1). Each is defined by the distribution of hydrogen bonds (donor or acceptor) on the purine and pyrimidine rings. Natural oligonucleotides use only two of these schemes, and one incompletely, as adenine (instead of diaminopurine) is used as the complement for uracil or thymine in natural nucleic acids.



R = β -ribofuranosyl or deoxyribofuranosyl

FIG. 1 Independently replicatable base pairs. Base pairs with different, and mutually exclusive, hydrogen-bonding patterns capable of forming Watson–Crick base pairs with three hydrogen bonds. The bases discussed in detail here are κ , X and an analogue of X designated π (shown in Fig. 2b) with the same hydrogen bonding pattern as X, but bases on a pyrazolopyrimidine ring system instead of a purine ring system (see text).

To obtain the desired hydrogen bonding, several of the 'unnatural' base pairs require a carbon-carbon bond joining the ribose ring with the heterocyclic base. Such structures are undoubtedly more difficult to prepare under prebiotic conditions than N-glycosides, perhaps explaining their absence from the repertoire of natural nucleotides. But such molecules can be obtained by synthetic chemistry which is not subject to prebiotic constraints.

Experiments with *iso*-C and *iso*-G (Fig. 1) have shown that bases containing novel hydrogen-bonding schemes can be incorporated by DNA and RNA polymerases¹⁵. Chemical considerations, however, suggest that *iso*-C might hydrolyse slowly to U under conditions of DNA synthesis, and that *iso*-G might exist to some extent in a minor tautomeric form complementary to U. Both of these problems were expected to complicate the selective incorporation of the *iso*-C/*iso*-G base pair in oligonucleotides also containing A and T, and complications were indeed observed¹⁵. Therefore, attention was directed towards developing base pairs (Fig. 1) for which both hydrolysis and tautomeric equilibria might be less problematic.

Selecting new bases κ , X and π

The pyrimidine 3- β -D-ribofuranosyl-(2,6-diaminopyrimidine), trivially designated as κ , presents a donor-acceptor-donor hydrogen bonding pattern to a complementary strand in a duplex structure. The pyrimidine κ was chosen after much work with the analogous pyridine base, 3- β -D-ribofuranosyl-(2,6-diaminopyridine)¹⁷. We have been unable so far to develop the oligonucleotide chemistry of this pyridine, perhaps because of the ease with which the heterocyclic ring is oxidized. The heterocyclic ring system of κ incorporates an additional nitrogen, rendering the ring more electron deficient than the pyridine system. κ as a deoxyribose derivative suitable for automated DNA synthesis was readily synthesized from a known precursor by the route shown in Fig. 2a. Crystals of the hydro-

TABLE 1 Melting temperatures for duplex formation

M	N	T_m [°C]	M	N	T_m [°C]
A	T	40	K	A	22
T	A	37	P	T	29
G	C	38	P	C	19
C	G	42	G	T	21
P	K	35	T	G	26
K	G	18			

The melting temperature T_m for duplexes formed with the general sequence: 5'-CAA $\overline{\text{M}}$ AAAG-3' and 3'-GTTT $\overline{\text{N}}$ TTTC-5', where $\overline{\text{M}}$ and $\overline{\text{N}}$ denote bases shown in Table. The melting curves were measured using a Perkin Elmer Lambda 2 UV/VIS spectrophotometer equipped with a Digital Controller thermoelectric temperature programmer. Readings at wavelength $\lambda = 260$ nm were taken with a temperature scan rate of $0.5^\circ\text{C min}^{-1}$. Profiles of absorbance versus temperature determined at slower heating rates or by increasing the temperature at 1°C min^{-1} gave identical results. Duplicate measurements were taken and found to yield identical T_m values to within $\pm 0.5^\circ\text{C}$. The cell dimensions were $1.2 \times 0.6 \times 1.9$ cm, with a path length of 0.1 cm. The buffer in all cases contained 1 M NaCl, 10 mM sodium phosphate and 0.1 mM EDTA in H_2O at pH 7. The concentration of each oligonucleotide was 50 μM .

chloride salt were grown from methanol, and a structure obtained by X-ray diffraction will be reported elsewhere.

Two purine analogues were chosen to complement κ . The first, xanthosine (X), is a natural base available commercially as both the nucleoside and nucleoside triphosphate. However, because of concerns that deoxyxanthosine might undergo depurination, another complementary base, 3- β -D-ribofuranosyl-(1-methyl-pyrazolo[4,3-d]pyrimidine-5,7(4H,6H)-dione), also known as 7-methyl oxoformycin B, and trivially designated here as π , was prepared by the route shown in Fig. 2b. In π , the heterocyclic base is joined to the pentose ring by a carbon-carbon bond. A crystal structure of π will also be reported elsewhere.

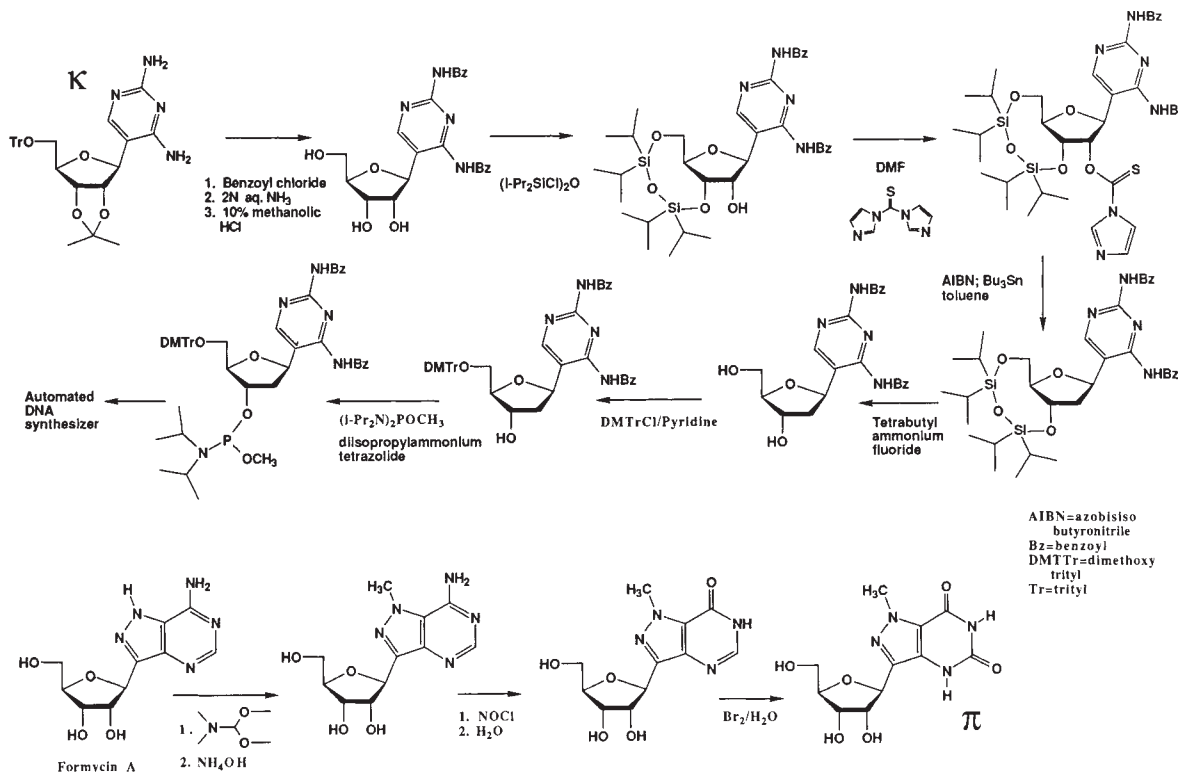


FIG. 2 Synthesis of derivatives of 3- β -D-ribofuranosyl-(2,6-diaminopyrimidine) (κ) and 3- β -D-ribofuranosyl-(1-methyl-pyrazolo[4,3-d]pyrimidine-5,7(4H,6H)-dione) (π). a, 3-(2,3-O-isopropylidene-5-O-trityl- β -D-ribofuranosyl)-(2,6-diaminopyrimidine), prepared by a six-step synthesis²⁴, was deoxygenated and then converted to a phosphoramidite suitable as a

building block for the automated chemical synthesis of DNA (^{31}P NMR doublet at 149.08 and 149.70 p.p.m. downfield from phosphoric acid)²⁵. b, Formycin A (Calbiochem, as monohydrate) was converted to 3- β -D-ribofuranosyl-(1-methyl-pyrazolo[4,3-d]pyrimidine-5,7(4H,6H)-dione) (π) by methylation and oxidation with nitrosyl chloride and bromine²⁶.

The κ - π base pair

Physical and chemical characterization of these bases was necessary to evaluate their suitability as replicatable components of a genetic alphabet. Although π exists in the crystal in the desired *anti* conformation, crystallographic studies of the hydrochloride salt of κ found the base in a *syn* conformation (data not shown), which is undesirable for the formation of a Watson-Crick base pair where both nucleoside bases must be in an *anti* conformation). Proton NMR studies were thus undertaken to determine whether κ could form a Watson-Crick base pair with π in solution.

In a solution of a derivative of κ in chloroform, a strong nuclear Overhauser enhancement (NOE) between the proton at C1' of the ribose ring and the proton at C4 of the heterocyclic ring (Fig. 3) indicated that κ adopts the undesired *syn* conformation when alone in solution. On addition of a protected derivative of the complementary purine nucleoside π , however, this NOE largely disappeared. Furthermore, when both κ and its complement are present, the resonances assigned to the amine protons of κ shift strongly downfield, as does the resonance of π assigned to the nitrogen flanked by the carbonyl groups. These facts together show that κ and π form a standard Watson-Crick base pair in a solution of chloroform.

Stability of duplexes containing the new base pair

To determine the effect of a κ - π on the stability of a DNA duplex, it was introduced synthetically into several oligonucleotides (Applied Biosystems, Table 1). Melting studies showed that duplexes containing a κ - π base pair are only slightly less stable than duplexes containing only natural bases (Table 1). Moreover, duplexes containing the new base pair are considerably more stable than those containing mismatches involving the new bases, which in turn have melting temperatures similar to duplexes containing mismatches of natural bases. The stabilities of various mismatches are consistent with the presumed stability of 'wobble' base pairs, which should be particularly important for the GT and A κ mismatches. These results

indicated that enzymatic incorporation of a new base selectively opposite its complement in a DNA template would be possible, provided that natural DNA and RNA polymerases accepted the new bases.

Enzymatic incorporation of a new base into RNA

For enzymatic studies with T7 RNA polymerase¹⁸, κ was complemented by the commercially available xanthosine triphosphate (XTP) instead of π , as xanthosine retains the essential hydrogen-bonding pattern of π , but lacks the methylated pyrazolo ring system that might be considered foreign by the polymerase. A promoter-template including a promoter sequence for T7 RNA polymerase (17 bases) followed by a short oligonucleotide segment (7 bases), the new base κ , 1 additional base, and a final A was synthesized, together with a complementary 18-base primer (Fig. 4). Incorporation of κ into the synthetic DNA templates was verified by digestion of samples of the

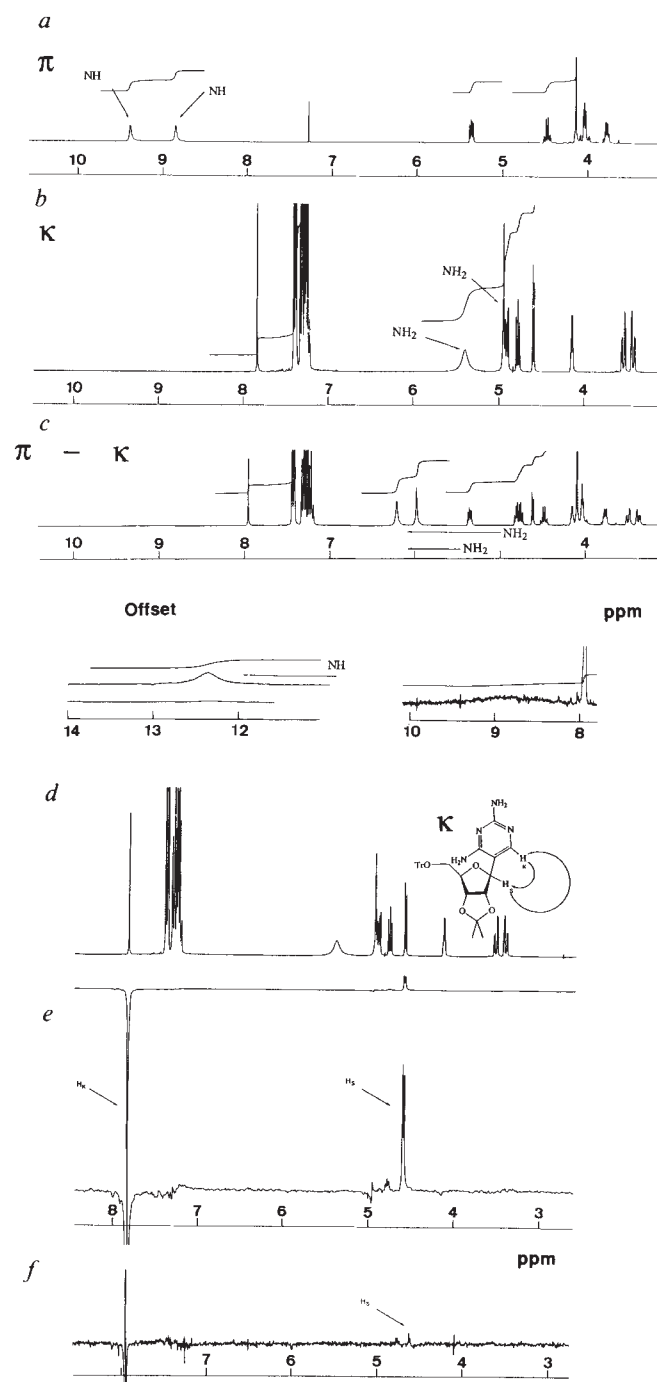


FIG. 3 NMR studies of the κ - π base pair in chloroform. The hydroxyl groups of the bases were protected to increase the solubility of the compounds in chloroform and to prevent the hydroxyl groups of the ribose from forming hydrogen bonds with protons on the base. All spectra were recorded at 24 ± 0.01 °C in chloroform (passed through basic alumina immediately before use and kept under nitrogen) on a 400 MHz Bruker NMR spectrophotometer. *a*, Spectrum of a derivative of κ (0.2M), with resonances assigned to the exchangeable hydrogens as indicated. *b*, Spectrum of the derivative of π indicated (0.2M), with resonances assigned to the exchangeable hydrogens indicated. *c*, Spectrum recorded with a 1:1 mixture of the derivatives of κ and π (0.2M each) showing downfield shift and broadening of the signals assigned to protons involved in the formation of hydrogen bonds (note offset at bottom). In the mixture, shifts greater than 0.8 p.p.m. are observed for the signals of both of the NH_2 protons of κ and one of the amide protons of π (presumably the 6-NH) is shifted by approximately 3 p.p.m. The other amide proton of π is broadened but its chemical shift remains largely unchanged. These shifts indicate a standard Watson-Crick base pair between the two groups. *d*, Spectrum of the derivative of κ (0.2M) in the *syn* conformation, indicating atoms expected to be close in space. *e*, Spectrum showing nuclear Overhauser enhancement (NOE) of the signal arising from the proton at C-1' of the ribose ring of κ following irradiation at the frequency of the resonance of the signal assigned to the proton at C-6 of the pyrimidine ring. The large NOE between H-6 and H-1' indicates that a substantial fraction of κ exists in the *syn* conformation at equilibrium, perhaps due to the formation of an intramolecular hydrogen bond between the 4-amino group and the furanose oxygen. *f*, Spectrum showing the NOE of the signal arising from the proton at C-1' of the ribose ring following irradiation at the frequency of the resonance of the signal assigned to the proton at C-6 of the pyrimidine ring, when κ is in the presence of an equal concentration of π . In the 1:1 mixture (0.2M of each nucleoside derivative), the NOE is significantly reduced, suggesting that the average distance between H-4 and H-1' in the base pair is considerably farther than in κ alone. This indicates that in the base pair with π , κ adopts an *anti* conformation.

template with snake venom phosphodiesterase, hydrolysis of the phosphate from the products with bacterial alkaline phosphatase, and analysis of the resulting nucleosides by HPLC (data not shown). Control templates containing T replacing κ were also prepared by synthesis. Transcription of the primed templates could be detected most simply by the incorporation of radiolabelled UMP (from α -labelled UTP) into a product RNA molecule 10 bases long (the 'full-length product').

When synthetic template 1 was incubated with labelled UTP and various other nucleoside triphosphates, full-length products

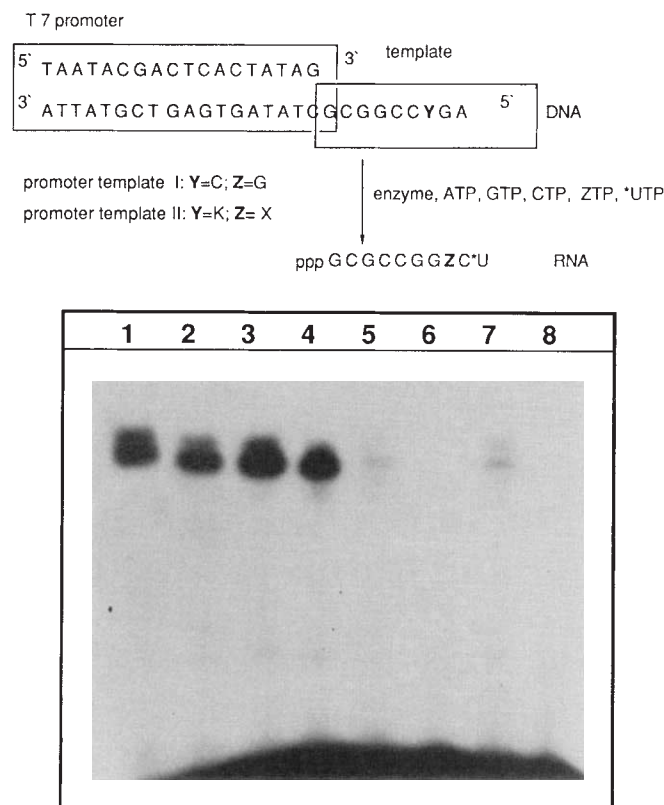


FIG. 4 Runoff transcription with T7 RNA polymerase. The procedure for using synthetic DNA templates to direct the synthesis of short RNA fragments by T7 polymerase was adopted from the literature¹⁷. Concentrations of single-stranded oligonucleotides were estimated by assuming an average molar extinction coefficient per base of $\epsilon_{260} = 1 \times 10^4 \text{ M}^{-1}$. Double-stranded templates were prepared by heating a 1:1 mixture of complementary single strands in triethyl ammonium bicarbonate (10 mM, pH 7) to 75 °C for 3 min, adding SSC buffer (6 M NaCl, 0.6 M sodium citrate, pH 7, diluted 20-fold), and cooling the solutions slowly to room temperature. They were stored at -20 °C. Standard transcription reactions contained template ($4.0 \mu\text{g ml}^{-1}$), T7 RNA polymerase (Pharmacia, 140 I.U.), nucleoside triphosphates (Pharmacia, 2.5 mM each), 20 mM MgCl_2 , spermidine (Fluka, 1 mM), dithiothreitol (Fluka, 5 mM), Tris-HCl buffer (Sigma, 40 mM, pH 8.1), bovine serum albumin (Boehringer Mannheim, $100 \mu\text{g ml}^{-1}$), RNase inhibitor (Boehringer Mannheim, 50 units), Triton X-100 (Sigma, 0.1%), and radiolabelled nucleotide [α - ^{32}P]UTP (Amersham, 3,000 C, mmol^{-1} , 1 μCi) in a total volume of 20 μl . Following the transcription reactions (4 h, 37 °C), the mixture was heated in loading buffer containing 7 M urea and dyes (0.02% bromophenol blue, 0.02% xylene cyanol, Eastman) for 5 min at 70–75 °C, and loaded immediately on an analytical denaturing 20% polyacrylamide gel. The products were separated on a 48 cm (0.6 mm thick) vertical gel apparatus, and the bands located by autoradiography. To determine yields, the product bands were excised, extracted with Protosol (300 μl each, 4 h, 35 °C) and counted in the presence of 2 drops of acetic acid and 9 ml of scintillant with a BETAmatic Kontron Liquid Scintillation counter. Lanes are designated (promoter-template; nucleoside triphosphates present). Lane 1 (I; GTP, CTP, [α - ^{32}P]UTP); lane 2 (II; GTP, CTP, [α - ^{32}P]UTP, XTP); lane 3 (II; GTP, CTP, [α - ^{32}P]UTP, 2 x XTP); lane 4 (II; GTP, CTP, [α - ^{32}P]UTP, 5 x XTP); lane 5 (II; GTP, CTP, [α - ^{32}P]UTP); lane 6 (I; CTP, [α - ^{32}P]UTP); lane 7 (II; GTP, CTP, [α - ^{32}P]UTP, ATP); lane 8 (II; CTP, [α - ^{32}P]UTP, XTP).

were observed in the presence of XTP (Fig. 4). The efficiency of synthesis of full-length product from templates with and without κ was approximately the same, provided that the necessary complementary nucleoside triphosphates were all present in the incubation mixtures. In the absence of XTP, a significant amount of full-length product could be detected only in the presence of ATP, and this at somewhat low concentrations (~24%, measured by scintillation counting of bands cut from the gel). Such a misincorporation presumably occurs through wobble base pairing, and is not infrequent even with natural bases when incorporation experiments are run in incubation mixtures that are missing one component¹⁹.

To determine whether misincorporation of A was diminished by competition with X, experiments were performed with tritiated XTP (synthesized from tritiated GTP by Demijanov oxidation)²⁰ and γ - ^{32}P -labelled GTP (which is incorporated, with the triphosphate intact, at the 5' end of the RNA product) together in an incubation mixture varying ratios. Full-length products from an incubation containing a 1:1 molar ratio of [^3H]XTP and unlabelled ATP were isolated by gel electrophoresis, the bands excised, and the relative amounts of ^3H and ^{32}P determined by scintillation counting. After correction for the specific activities of the starting nucleotides, the misincorporation of adenosine into the product at a XTP:ATP ratio of 1:1 was reduced to ~14%. Infidelity further decreases with increasing ratios of X:A, and probably stems from errors in the action of the polymerase itself¹⁹, rather than from the presence of minor tautomers of the bases.

Incorporation of the κ -X base pair into DNA

DNA polymerases frequently display higher fidelity than RNA polymerases. To assess the potential for incorporating the κ -X base pair in a replicatable system, the ability of DNA polymerase to incorporate X in a growing DNA product opposite κ in a DNA template was examined. A set of primer-templates (Fig. 5) were prepared containing either κ , C or T (the latter two serving as control templates). Incorporation of κ into the synthetic DNA templates was again verified by digestion of samples of the template with snake venom phosphodiesterase, removal of the phosphate from the products by bacterial alkaline phosphatase, and analysis of the resulting nucleosides by HPLC (data not shown). As before, the last base in the template was a unique A, permitting the detection of full length products most simply by autoradiography following the incorporation of [α - ^{32}P]TTP.

The synthetic primer templates were incubated with the Klenow fragment of DNA polymerase I (*Pol* I)²¹ in the presence of various nucleoside triphosphates, and the products analysed by gel electrophoresis (Fig. 5). κ in the template directed the incorporation of XTP into full-length product. On electrophoresis, the product containing X migrated faster than the analogous products containing G or A, presumably because the xanthine heterocycle carries an additional negative charge under the conditions of the electrophoresis due to its low pK_a of 5.7. Direct evidence for the incorporation of xanthosine was obtained by digestion of the product oligonucleotide, treatment with kinase and electrophoretic analysis.

To measure the relative efficiency as templates of the oligonucleotides containing different bases, product bands from electrophoresis gels were excised and their radioactivity determined by liquid scintillation counting. Templates containing κ were ~70% as efficient at directing the synthesis of full-length product (in the presence of XTP) as those containing only natural bases.

The fidelity of incorporation of X opposite κ was examined by incubating templates containing C, T and κ with purine triphosphates separately and in competition (Fig. 5). As expected, the fidelity of incorporation was considerably higher with DNA polymerase than with T7 RNA polymerase. Essentially no G or A was incorporated by the Klenow fragment of DNA polymerase opposite κ , and essentially no X was incorporated

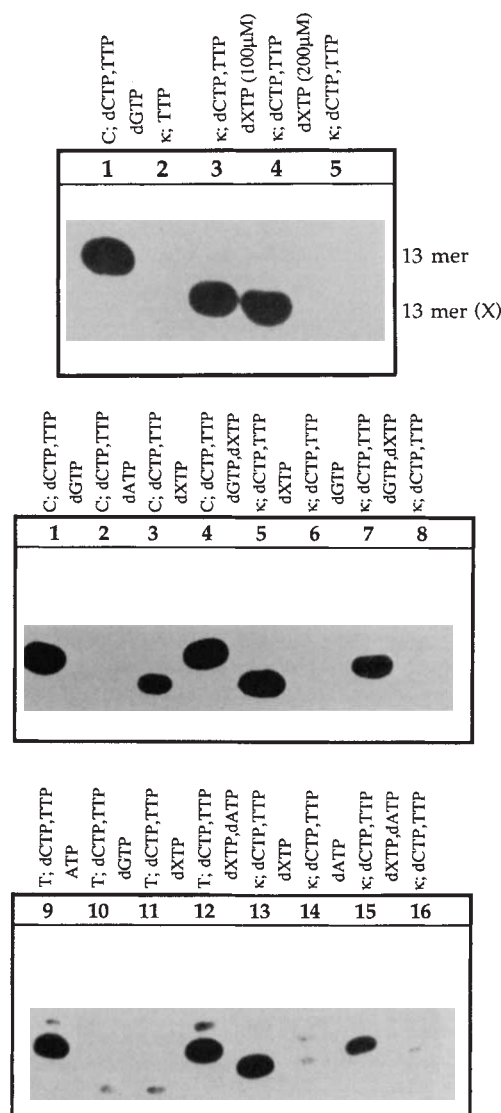
FIG. 5 Incorporation experiments with DNA polymerase I (Klenow fragment). The standard fill-in reactions contained Tris-HCl buffer (Sigma, 50 mM, pH 7.5), MgCl₂ (10 mM), bovine serum albumin (Boehringer Mannheim, 0.10 µg), DTT (Fluka, 1 mM), nucleoside triphosphates (Pharmacia, 100 µM) as specified below, [α -³²P]TTP (Amersham, 3,000 Ci/mmol⁻¹, 1 µCi) and Klenow enzyme (Boehringer Mannheim, 1 unit) in a total volume of 20 µl (ref. 20). The mixtures were incubated at 23 °C for 15 min, and then heated in loading buffer containing 7 M urea and dyes (0.02% bromophenol blue and xylene cyanol, Eastman) for 3 min at 70–75 °C, and loaded immediately on an analytical 20% denaturing polyacrylamide gel. The products were separated on a 48 cm (0.6 mm thick) vertical gel apparatus, and the product bands located by autoradiography. To determine yields, the product bands were excised, extracted with Protosol (300 µl each, 4 h, 35 °C) and counted in the presence of 2 drops of acetic acid and 10 ml of scintillant with a BetaMatic Kontron Instrument. The template-primers used had the general formula, 5'-GATTTTGA-3', and 3'-CTAAACTGGNGA-5', where for primer-template I, N = κ ; for primer-template II, N = C; and for primer-template III, N = T. Lanes are designated (primer-template; nucleoside triphosphates present) Top gel: lane 1 (II; dCTP, [α -³²P]TTP, dGTP); lane 2 (I; [α -³²P]TTP); lane 3 (I; dCTP, [α -³²P]TTP, dXTP (100 µM)); lane 4 (I; dCTP, [α -³²P]TTP, dXTP (200 µM)); lane 5 (I; dCTP, [α -³²P]TTP). Middle gel: lane 1 (II; dCTP, [α -³²P]TTP, dGTP); lane 2 (II; dCTP, [α -³²P]TTP, dATP); lane 3 (II; dCTP, [α -³²P]TTP, dXTP); lane 4 (II; dCTP, [α -³²P]TTP, dGTP, dXTP); lane 5 (I; dCTP, [α -³²P]TTP, dXTP); lane 6 (I; dCTP, [α -³²P]TTP, dGTP); lane 7 (I; dCTP, [α -³²P]TTP, dGTP, dXTP); lane 8 (I; dCTP, [α -³²P]TTP). Bottom gel: lane 9 (III; dCTP, [α -³²P]TTP, dATP); lane 10 (III; dCTP, [α -³²P]TTP, dGTP); lane 11 (III; dCTP, [α -³²P]TTP, dXTP); lane 12 (III; dCTP, [α -³²P]TTP, dATP); lane 13 (I; dCTP, [α -³²P]TTP, dXTP); lane 14 (I; dCTP, [α -³²P]TTP, dATP); lane 15 (I; dCTP, [α -³²P]TTP, dXTP, dATP); lane 16 (I; dCTP, [α -³²P]TTP). The middle and bottom gels are overexposed to allow weak bands (which possibly indicate infidelity) to show clearly. Oligonucleotide products containing X run faster than corresponding products containing A or G, presumably because the xanthine heterocycle ($pK_a = 5.7$) is deprotonated under the conditions of the electrophoresis. The gel 'smiles' significantly; based on the position of the markers, the main band in lane 16 comigrates with other oligonucleotides on the gel containing xanthosine.

opposite T in the template. The only evidence of infidelity was a low level (~5%) of X misincorporated opposite C in the template when GTP was missing from the incubation mixture. This misincorporation was not observed at all when GTP and XTP were present in a 1:1 ratio.

Discussion

These studies show that although it is present in the crystal and in dilute solutions of chloroform in the *syn* conformation, κ will form a base pair with an appropriate purine nucleoside in organic solvents. Further, templates containing κ direct both T7 RNA polymerase and the Klenow fragment of DNA polymerase 1 to incorporate xanthosine as the complementary purine nucleoside into DNA and RNA products. With DNA polymerase, the incorporation of the new base pair proceeds with high fidelity. Thus, this work demonstrates the feasibility of expanding the genetic alphabet, increasing the number of letters that can be enzymatically incorporated into oligonucleotides by template-directed polymerization.

This expanded genetic alphabet should allow more diversity in the functional groups available to RNA, as synthetic routes



to many functionalized bases are described in the literature, and many pyrimidines functionalized at the 5-position are accepted by polymerases²². Incorporation of such functionalized bases into RNA should provide RNA molecules with the potential for greatly increased catalytic power, including perhaps RNA molecules that catalyse their own replication.

The new base pair described here should also find use in studies of the structure of biologically important RNA and DNA molecules²³ and protein-nucleic acid interactions. More speculatively, the extra letters in the nucleoside alphabet might eventually be used to expand the genetic code, increasing the number of amino acids that can be incorporated translationally into proteins²⁴. □

Received 11 September; accepted 14 November 1989.

- Gilbert, W. *Nature* **319**, 818 (1986).
- Orgel, L. E. *J. theor. Biol.* **123**, 127–149 (1986).
- White, H. B. III *J. molec. Evol.* **7**, 101–104 (1976).
- Visser, C. M. & Kelllogg, R. M. *J. molec. Evol.* **11**, 171–187 (1978).
- Rich, A. in *Horizons in Biochemistry* 103 (eds Kasha, M. & Pullmann, B.) 103–126 (Academic, New York, 1962).
- Schram, G., Götsch, H. & Pollmann, W. *Angew. Chem.* **73**, 619 (1961).
- Usher, D. A. & McHale, A. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1149–1153 (1976).
- Cech, T. R., Zaig, A. J. & Grabowski, P. J. *Cell* **27**, 487–496 (1981).
- Guerrier-Takada, C. & Altman, S. *Science* **223**, 285–286 (1983).
- Doudna, J. A. & Szostak, J. W. *Nature* **339**, 519–522 (1989).
- Joyce, G. F., Schwartz, A. W., Miller, S. L. & Orgel, L. E. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4398–4402 (1987).
- Weiner, A. M. & Maizels, N. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7383–7387 (1987).
- Benner, S. A., Ellington, A. D. & Tauer, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7054–7058 (1989).
- Benner, S. A. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **52**, 53–63 (1987).

- Switzer, C. Y., Moroney, S. E. & Benner, S. A. *J. Am. Chem. Soc.* **111**, 8322–8323 (1989).
- Uhlenbeck, O. C. *Nature* **328**, 596–600 (1987).
- Picirilli, J. A. thesis, Harvard Univ. (1989).
- Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. *Nucleic Acids Res.* **15**, 8783–8798 (1987).
- Coleman, J. E., Martin, C. T. & Muller, D. K. *Biochemistry* **27**, 3966–3974 (1988).
- Roy, K. B. & Miles, H. T. *Nucleosides Nucleotides* **2**, 231–242 (1983).
- Cobianchi, F. & Wilson, S. H. *Meth. Enzym.* **152**, 94–110 (1987).
- Leary, J. L., Brigati, D. J. & Ward, D. C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4045–4049 (1983).
- Chen, T. R., Churchill, M. E. A., Tullius, T. D., Kallenbach, N. R. & Seemann, N. C. *Biochemistry* **27**, 6032–6038 (1988).
- Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. *Science* **244**, 182–188 (1989).

ACKNOWLEDGEMENTS. We are indebted to Dr Christopher Y. Switzer for assistance with DNA polymerase, and to Dr B. Schweizer for obtaining crystal structures of π and κ which will be described in detail elsewhere. We thank Sandoz AG, the Swiss National Science Foundation, the Eidgenössische Technische Hochschule, and the Verband der Deutschen Chemischen Industrie for support.

Redesign of the coenzyme specificity of a dehydrogenase by protein engineering

Nigel S. Scrutton, Alan Berry & Richard N. Perham*

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

Directed mutagenesis and molecular modelling have been used to identify a set of amino-acid side chains in glutathione reductase that confer specificity for the coenzyme NADP⁺. Systematic replacement of these amino acids, all of which occur in a 'fingerprint' structural motif in the NADP⁺-binding domain, leaves the substrate specificity unchanged but converts the enzyme into one displaying a marked preference for the coenzyme NAD⁺.

THERE are two main classes of redox coenzyme found in biological chemistry, the nicotinamide nucleotides and the flavins, both of which act as ancillary electron carriers for a wide variety of enzymes involved in metabolism¹. The degradation of fuel molecules is an oxidative exergonic process yielding energy in the form of ATP, and in aerobic organisms is coupled to the reduction of O₂. Biosynthesis, by contrast, is a reductive endergonic process, generally requiring ATP and a source of reducing power.

Nicotinamide adenine dinucleotide (NAD⁺) is a ubiquitous coenzyme in biological oxidations. In the presence of the substrate and the relevant enzyme it undergoes reduction, its nicotinamide ring accepting a proton and two electrons (equivalent to a hydride ion) to yield the dihydropyridine form (NADH). It is the subsequent oxidation of NADH by the cytochromes of the electron transport chain that constitutes the principal source of ATP in aerobic organisms. Another similar and ubiquitous coenzyme, NADP⁺, differs structurally from NAD⁺ only by the presence of an additional phosphate group esterified to the 2'-hydroxyl group of its AMP moiety. NAD⁺ is used, however, almost exclusively in the oxidative degradations that yield ATP, whereas NADPH is confined, with few exceptions, to the reactions of reductive biosynthesis. It is the possession of the extra phosphate group by NADP⁺ that permits the biosynthetic enzymes to make this distinction, a simple yet impressive example of the power of molecular recognition in biochemical reactions.

The methods of protein engineering and site-directed mutagenesis provide the technology required to study recognition processes of this kind. Recently, much effort has been made to study substrate specificity in enzymes, and several attempts have been made rationally to change the specificity of particular enzymes²⁻⁵. In one case at least, significant success has been achieved; that is, in the conversion of a lactate dehydrogenase to a malate dehydrogenase⁶. Another obvious and more widely applicable challenge is to switch the coenzyme specificity of an enzyme, for example, from using NADP⁺ to NAD⁺. Apart from its intrinsic interest in the study of molecular recognition, such a switch could permit a better understanding of the different roles of the coenzymes in catabolic and anabolic processes. Also, it is likely to prove of importance in biotechnology, given that NADP⁺ is much more expensive than NAD⁺.

The prerequisites for redesigning a protein in this way are a

detailed knowledge of the structural features in the active site that impart the specificity for the coenzyme, generally obtained from a crystal structure of the enzyme-coenzyme complex, and a cloned and expressed structural gene encoding the enzyme. These prerequisites are favourably met in the enzyme glutathione reductase (EC 1.6.4.2.). Glutathione reductase catalyses the reduction of oxidized glutathione (GSSG) to glutathione (GSH) using NADPH as reducing coenzyme. The enzyme is of widespread importance in maintaining the reduced state of the cell and in more specialized pathways, such as DNA biosynthesis^{7,8}. With only one exception—the enzyme from *Chromatium vinosum*⁹—glutathione reductase requires NADPH as a reducing coenzyme. An excellent three-dimensional structure of the enzyme from human erythrocytes has been determined^{10,11} and recently refined to 0.15-nm resolution¹². The structural gene (*gor*) for the enzyme from *Escherichia coli* has been cloned and sequenced¹³ and overexpressed^{14,15}. Also, the structure of the human enzyme is a good model for protein engineering experiments with the *E. coli gor* gene^{16,17}.

The NADP⁺-binding site

Glutathione reductase is a member of a highly homologous family of enzymes, the flavoprotein disulphide oxidoreductases^{18,19}, most of which function with NADPH as a coenzyme. But one member of the family, dihydrolipoamide dehydrogenase (EC 1.6.4.1), uses NAD(H), and, although a high-resolution crystal structure for this enzyme has not yet been determined^{20,21}, the amino-acid sequence of the *E. coli* dihydrolipoamide dehydrogenase is sufficiently similar to that of the human glutathione reductase for the three-dimensional structure of the latter enzyme to act as a plausible model for both²². So there is a guide of how NAD⁺ is likely to be bound in an enzyme which is closely related to one that has a great preference for NADP⁺.

In human glutathione reductase, NADPH is bound in an extended conformation (Fig. 1a) in a cleft that constitutes part of the active site^{10,11,23}. Within the structure of the enzyme are several clearly delineated functional domains, one of which is primarily responsible for binding the NADPH and comprises the characteristic βαβαβ dinucleotide-binding fold typical of NAD⁺-linked dehydrogenases²⁴. The nicotinamide moiety stacks with the middle ring of the enzyme-bound flavin, where it makes additional contacts with amino-acid side chains contributed by the FAD-binding domain, the central domain and the subunit interface domain²³. The pyrophosphate bridge of the coenzyme is bound at the C-terminal end of a parallel β-sheet in the NADP-binding domain²³, its negative charge stabilized by the dipole at the N-terminal end of an α-helix beginning at residue Gly 196, another structural feature typical of this dinucleotide-binding motif²⁴⁻²⁶.

NADH also binds to human glutathione reductase, but with a Michaelis constant (*K_m*) ~60-times higher than that for NADPH, in part perhaps because it lacks the 2'-phosphate group of the latter²³. The binding site for the 2'-phosphate includes three prominent positively charged residues—Arg 218, His 219 and Arg 224. Of these, the two arginine residues make several short contacts, whereas His 219 is a little removed (Fig. 1a, b). This is consistent with the phosphate group carrying two negative charges²³, as it does in the nonhomologous enzyme dihydrofolate reductase²⁷. NADH binds to human glutathione reductase in a similar way to NADPH, except that an inorganic

* To whom correspondence should be addressed.

phosphate ion substitutes for the missing 2'-phosphate group on the coenzyme²³.

All the amino-acid residues identified as contributing to the active site or otherwise participating in the catalytic mechanism in human glutathione reductase^{11,23} are conserved in the *E. coli* enzyme, with the exception of His 219¹³, which in the *E. coli* enzyme is replaced by a lysine residue. Given that this residue may be involved in binding the 2'-phosphate group of NADP⁺, this substitution can be viewed as a conservative one. The only other main difference in the primary structure of the human and *E. coli* glutathione reductases is the presence in the former of ~18 additional residues at the N terminus. These are not visible in the X-ray crystallographic structure of the human enzyme and are believed to constitute a disordered or flexible segment of polypeptide chain unlikely to contribute to the mechanism in any way¹⁰⁻¹³. It means, however, that there is a difference in residue numbering between the human and *E. coli* enzymes. We now describe a series of mutations in the *E. coli gor* gene that positively identifies the amino-acid side chains that confer the specificity for NADPH in glutathione reductase and discriminate against NADH, and which has enabled us to change the enzyme from one that preferentially uses NADPH to one that has a preference for NADH.

Arg 198 and Arg 204 confer specificity for NADPH

Arg 218 and Arg 224 of human glutathione reductase are both retained in virtually all the NADP⁺-requiring enzymes of the flavoprotein oxidoreductase family, but not in the NAD⁺-requiring dihydrolipoamide dehydrogenase (Table 1). His 219 is much less well conserved, having mutated to a lysine residue in the *E. coli* glutathione reductase and to uncharged serine or asparagine residues in the mercuric reductases and trypanothione reductase (Table 1). So the two arginine residues seem to be of particular importance in binding the 2'-phosphate group of NADPH.

Using site-directed mutagenesis on the *E. coli gor* gene, carried out as described elsewhere¹⁵⁻¹⁷, we converted Arg 198 (equivalent to Arg 218 in human glutathione reductase) to methionine (R198M), the residue that occupies the equivalent position in *E. coli* dihydrolipoamide dehydrogenase (Table 1). In choosing the product of the mutation, a long aliphatic side chain was also favoured by the extensive interaction that the side chain of Arg 218 in human glutathione reductase makes with the adenine ring of the bound NADPH (Fig. 1b)²³. Residue Arg 204 (equivalent to Arg 224 in the human enzyme) was initially converted to leucine (R204L). The residue at the equivalent position in *E. coli* dihydrolipoamide dehydrogenase is proline (Table 1) but, given the structural constraints imposed by proline residues²⁸, we reserved the proline exchange for later (see below).

The mutated *gor* genes (R198M, R204L and the double mutant R198MR204L) were expressed from plasmid pKGR4 under the control of the *tac* promoter in a *gor*-deletion strain of *E. coli*¹⁵.

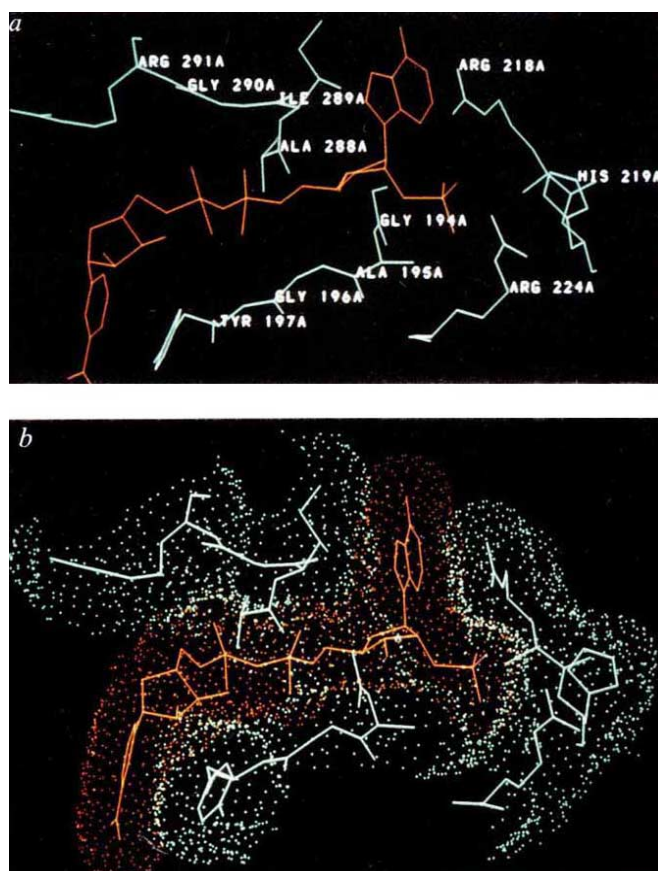
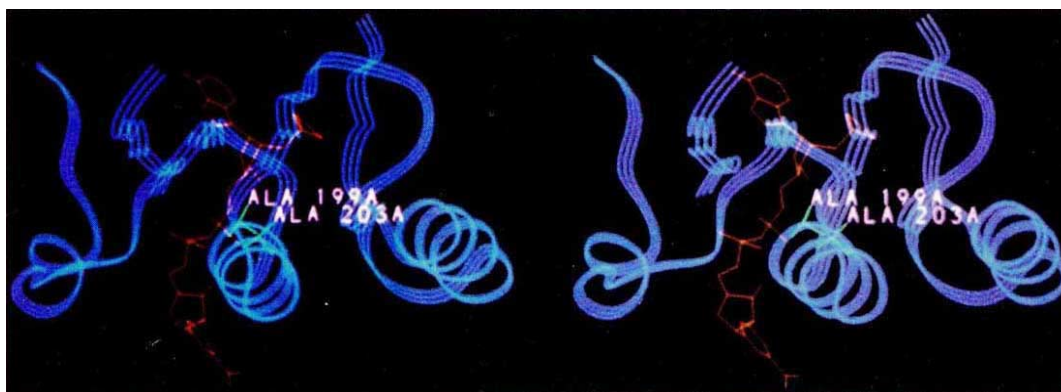


FIG. 1 Molecular-graphic representations of the NADPH-binding site of human glutathione reductase, produced from coordinates provided by Professor G. E. Schulz, using the programme FRODO on an Evans & Sutherland PS390 graphics terminal. *a*, Portions of polypeptide shown are Gly 194 to Tyr 197, Ala 288 to Arg 291 and the residues believed to be involved in binding the 2'-phosphate, Arg 218, His 219 and Arg 224. *b*, Same portions of polypeptide as shown in *a* together with a water-accessible surface superimposed around the polypeptide and nicotinamide cofactor.

This increases the expression of the *gor* gene to ~40,000-times that of wild-type *E. coli* and permits a simple purification of the enzyme to homogeneity¹⁵. When assayed for catalytic activity with NADPH, the optimum pH values for the mutant enzymes were lower than that for the wild-type enzyme (Table 2A). This is consistent with a need to protonate the 2'-phosphate of the NADPH to compensate for the loss of positively charged side chains from the 2'-phosphate-binding site in the enzyme.

The values of the apparent K_m for NADPH and for the catalytic constant (turnover number) k_{cat} were determined for each enzyme under conditions of optimal pH and at a saturating

FIG. 2 Molecular graphics representation (stereo-pair) of the α -helix containing the residues Ala 199 and Ala 203 in human glutathione reductase. The methyl side chains of the two alanine residues are shown interacting with the β -strands of the dinucleotide-binding domain.



GSSG concentration of 1.2 mM. Substitution of either arginine residue caused a large rise (~ 25 -fold) in K_m and a modest fall in k_{cat} ; substitution of both arginine residues caused no further change in K_m , but there was a further fall in k_{cat} to $\sim 5\%$ of the wild-type value (Table 2A). The effects of substituting the two arginine residues in the 2'-phosphate-binding site therefore are a substantially decreased affinity of glutathione reductase for NADPH and, almost certainly, a catalytically less favourable configuration for the NADPH bound in the double-mutant enzyme.

We also assayed wild-type and mutant enzymes using NADH as the coenzyme. In all three cases, the optimal pH for the mutant enzymes was similar to that for the wild-type enzyme (Table 2D). Because NADH carries no 2'-phosphate group, this observation supports the hypothesis that it could be the need to maintain electrical neutrality in what was the 2'-phosphate-binding site that leads to a lowering of the optimal pH for the mutant enzymes when using NADPH as coenzyme. Estimation of the steady-state kinetic parameters, measured at the optimal pH in the presence of NADH, revealed no significant change in k_{cat} and only a 3- to 4-fold decrease in the value of K_m (Table 2D). Thus, although the positively charged residues Arg 198 and Arg 204 in *E. coli* glutathione reductase clearly make an important contribution to the binding of the NADPH molecules, and confer the observed coenzyme specificity, their presence does not greatly diminish the propensity, albeit weak, of the enzyme to bind NADH. Conversely, therefore, it is reasonable to infer that replacement of the two arginine residues will contribute to the rejection of NADPH as a coenzyme by dihydrolipoamide dehydrogenase, but will not predispose this enzyme to bind NADH. Some other structural feature must be sought to account for this.

The acquisition of NADH-dependent activity

Common to the NAD⁺-binding domains of many enzymes is a $\beta\alpha\beta$ -fold^{24,29}, centred around a highly conserved sequence Gly-

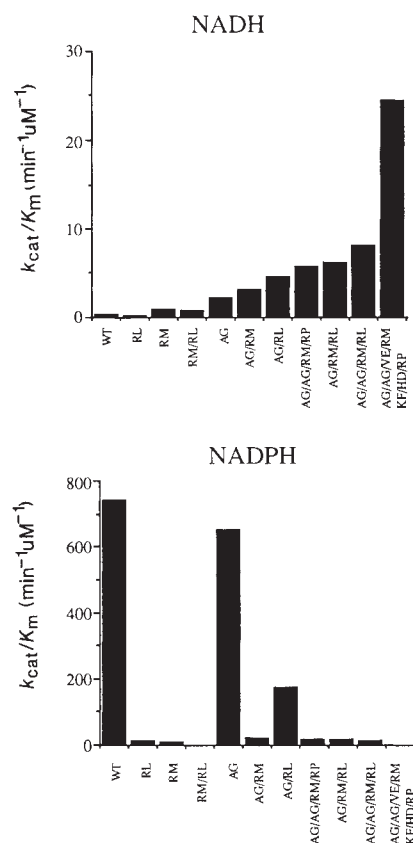


FIG. 3 Histograms illustrating the effects of mutation on k_{cat}/K_m for the two nicotinamide coenzymes. Values of k_{cat} and K_m are taken from Table 2. WT, wild type. Single-letter amino-acid code used to denote mutations.

TABLE 1 Alignment of sequences of flavoprotein disulphide oxidoreductases

NADP(H) binding																			
Glutathione reductase																			
<i>E. coli</i>	190	G	A	K	T	H	L	F	V	R	K	H	A	P	L	R	S	F	D
Human	210	G	S	K	T	S	L	M	I	R	H	D	K	V	L	R	S	F	D
<i>P. aeruginosa</i>	188	G	A	E	T	T	L	L	Y	R	R	D	L	F	L	R	G	F	D
Mercuric reductase																			
<i>S. aureus</i> pl258	273	G	T	E	V	T	L	M	Q	R	S	E	R	L	F	K	T	Y	D
<i>P. aeruginosa</i> Tn501	293	G	S	K	V	T	V	L	A	R	N	T	L	F	F	R	E	—	D
<i>S. flexnerii</i> R100	292	G	A	K	V	T	I	L	A	R	S	T	L	F	F	R	E	—	D
Trypanothione reductase																			
<i>T. congolense</i>	214	G	G	K	V	T	L	C	Y	R	N	N	P	I	L	R	G	F	D
NAD(H) binding																			
Dihydrolipoamide dehydrogenase																			
<i>E. coli</i>	196	G	S	Q	I	D	V	V	E	M	F	D	Q	V	I	P	A	A	D
<i>B. stearothermophilus</i>	199	G	T	K	V	T	I	L	E	G	A	G	E	I	L	S	G	F	E
Yeast	225	G	S	K	V	T	V	V	E	F	Q	P	Q	I	G	A	S	M	D
Human	236	G	A	D	V	T	A	V	E	F	L	G	H	V	G	G	V	G	I
Pig	236	G	A	D	V	T	A	V	E	L	L	G	H	V	G	G	I	G	I
<i>P. fluorescens</i>	204	G	A	E	V	T	V	L	E	A	L	D	K	F	L	P	A	A	D
<i>A. vinelandii</i>	203	G	A	E	V	T	V	L	E	A	M	D	K	F	L	P	A	V	D

Alignment of the primary sequences of flavoprotein disulphide oxidoreductases around the region of Arg 218 and Arg 224 of human glutathione reductase. The residues inferred to be involved in interacting with the 2'-phosphate of NADPH in enzymes using NADP(H) are illustrated in upright bold. The residues in equivalent positions in enzymes using NAD(H) are also shown in upright bold. In the NADH-dependent enzymes, the conserved glutamate residue involved in binding the 2'-OH of the ribose of NADH is shown in italicized bold, as are the residues at the equivalent positions in the NADP(H)-dependent enzymes. The numbers indicate the position in the primary structure of the first residues shown for each sequence. Specific references: glutathione reductase, *E. coli* (ref. 13); human (ref. 33); *Pseudomonas aeruginosa* (N. L. Brown, personal communication); mercuric reductase, *Staphylococcus aureus* (ref. 34); *P. aeruginosa* (ref. 35); *Shigella flexnerii* (ref. 36); trypanothione reductase, *Trypanosoma congolense* (ref. 37); dihydrolipoamide dehydrogenase, *E. coli* (ref. 38); *Bacillus stearothermophilus* (A. Borges, C. F. Hawkins and R.N.P., unpublished data); yeast (ref. 39); human and pig (ref. 40); *P. fluorescens* (ref. 41); *Azotobacter vinelandii* (ref. 42).

X-Gly-X-X-Gly (where X is any amino acid) that constitutes a tight turn at the end of the first strand of a β -sheet and marks the beginning of the succeeding α -helix²⁵. The first glycine residue is essential for the tightness of the turn, the second (at the N terminus of the helix) allows the dinucleotide to be bound without obstruction from an amino-acid side chain at this position, and the third seems to provide space for a close interaction between the β -strands and the α -helix. The helix dipole contributes significantly to the binding of the coenzyme by interacting favourably with the pyrophosphate moiety^{25,30}. Other conserved features of this fingerprint region are (1) a hydrophilic residue at the N terminus of the first β -strand (function unknown), (2) a hydrophobic core composed of six small residues, and (3) a negatively charged residue at the C terminus of the second β -strand which forms a hydrogen bond with the 2'-hydroxyl group of the adenine ribose of the NAD⁺ (refs 25 and 31). A compilation of all the amino-acid sequences around the conserved glycine residues available to us from the data bases is shown in Table 3.

The number of NADP⁺-binding domains for which the tertiary structure is known is much smaller and knowledge about them is less detailed, but in several instances (including glutathione reductase) the same fingerprint region can be identified, although with some significant differences^{25,26}. First, the negatively charged residue at the C-terminal end of the second

β -strand is replaced, presumably to accommodate the 2'-phosphate group on the coenzyme²⁵. This is clearly visible from Table 1, which lists the sequences in this region for the family of flavoprotein disulphide oxidoreductases. Secondly, the third glycine residue of the conserved trio is replaced by alanine^{25,26}, and there is almost always another alanine a further four residues along the helix towards the C terminus²⁶. This feature is illustrated in Table 3. The sequences of some NADPH-requiring enzymes do not fit this pattern (not shown in Table 3) and their NADPH-binding sites cannot incorporate the same compact $\beta\alpha\beta$ -fold^{25,26}.

Substitution of Ala 179 and Ala 183. These two alanine residues in *E. coli* glutathione reductase, characteristic of the NADP⁺-binding domain (Table 3), are located in the α -helix of the $\beta\alpha\beta$ -fold. The spacing of four amino acids between them means that the two alanine residues must lie on the same side of the helix. In this position, their methyl groups interdigitate between the neighbouring β -strands (Fig. 2), occupying more space than the corresponding glycine residues of the NAD⁺-binding fold and so modifying the local polypeptide conformation. This could be associated with a change of coenzyme specificity. To test this possibility, further mutants were created and characterized.

The effect of introducing a glycine residue at position 179 (A179G) was a dramatic fall of K_m for NADH to a value almost 40-fold lower than the K_m for the wild-type enzyme, both in

TABLE 2 Kinetic parameters of native and mutant enzymes

Kinetic parameters with NADPH as coenzyme					
Enzyme	K_m for NADPH (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\text{min}^{-1} \mu\text{M}^{-1}$)	pH (optimum)	k_{cat}/K_m (Mutant: Wild type)
A Wild type	21.7 $\pm$ 2.0	16,000 $\pm$ 350	740.0	7.5	1.00
R204L	450.0 $\pm$ 30	5,760 $\pm$ 130	12.8	6.2	0.017
R198M	570.0 $\pm$ 100	4,860 $\pm$ 370	8.5	5.4	0.011
RM/RL	480.0 $\pm$ 100	790 $\pm$ 65	1.6	5.0	0.002
B A179G	12.0 $\pm$ 0.5	7,800 $\pm$ 100	650.0	7.5	0.88
AG/RM	723.0 $\pm$ 44	15,600 $\pm$ 460	21.6	6.2	0.029
AG/RL	72.6 $\pm$ 10	12,500 $\pm$ 1,100	172.0	6.8	0.23
AG/RM/RL	27.0 $\pm$ 4	480 $\pm$ 30	18.0	5.0	0.024
AG/A183G/RM/RL	140.0 $\pm$ 30	1,800 $\pm$ 280	12.9	5.7	0.017
AG/AG/RM/R204P	80.3 $\pm$ 6	1,250 $\pm$ 60	15.6	5.4	0.021
C AG/AG/V197E/ RM/K199F/H200D/RP	220.0 $\pm$ 30	660 $\pm$ 75	3.0	5.4	0.004
Kinetic parameters with NADH as coenzyme					
Enzyme	K_m for NADH (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\text{min}^{-1} \mu\text{M}^{-1}$)	pH (optimum)	k_{cat}/K_m (Mutant: Wild type)
D Wild type	2,000.0 $\pm$ 480	680 $\pm$ 110	0.34	4.7	1.00
R204L	720.0 $\pm$ 120	120 $\pm$ 10	0.17	4.7	0.50
R198M	780.0 $\pm$ 130	750 $\pm$ 60	0.96	4.7	2.82
RM/RL	530.0 $\pm$ 80	420 $\pm$ 30	0.8	4.7	2.35
E A179G	51.0 $\pm$ 14	110 $\pm$ 14	2.16	4.7	6.35
AG/RM	64.0 $\pm$ 9	200 $\pm$ 15	3.13	4.7	9.20
AG/RL	46.0 $\pm$ 8	210 $\pm$ 20	4.57	4.7	13.4
AG/RM/RL	80.0 $\pm$ 9	490 $\pm$ 30	6.13	5.0	18.0
AG/A183G/RM/RL	67.0 $\pm$ 17	540 $\pm$ 90	8.06	5.4	23.7
AG/AG/RM/R204P	120.0 $\pm$ 25	690 $\pm$ 100	5.75	5.4	16.9
F AG/AG/V197E/ RM/K199F/H200D/RP	86.0 $\pm$ 11	2,100 $\pm$ 150	24.4	5.4	71.8

Kinetic parameters of native and mutant enzymes. Single-letter amino-acid code used to denote mutations. Mutant genes were generated by the phosphorothioate procedure of Taylor *et al.*⁴³ as described previously¹⁵, using the following mutagenic oligonucleotides and the template K19gor EcoRI 3'8 (ref. 15) unless otherwise stated: R198M, 5'-ATCTGTTTGTGATGAACATGCGCCG-3'; R204L, 5'-ATGCGCCGCTGCTGAGCTTCGACCCG-3'; A179G, 5'-GGTTACATCGGTGTTGAGCTGGCGGG-3'; A183G (A179G R198M R204L mutant template K19gor EcoRI 3'8), 5'-TGTTGAGCTGGGTGGCGTGATT-3'; R204P (A179G A183G R198M R204L mutant template K19gor EcoRI 3'8), 5'-ATGCGCCGCTGCCAGCTTCGACCCG-3'; V197E K199F H200D (A179G A183G R198M R204P mutant template K19gor EcoRI 3'8), 5'-ACGCATCTGTTGAAATGTTGATGCGCCGCTGC-3'. The genes were expressed from the plasmid pKK223-3 in *E. coli* strain NS3 and the mutant proteins were purified as described elsewhere¹⁵. In all cases the whole gene was resequenced to check that no spurious mutations had arisen during the mutagenesis procedure. The kinetic parameters were determined at 30 °C in a mixture of 0.1 M citric acid and 0.2 M sodium phosphate adjusted to the optimum pH value of each enzyme at a fixed concentration of oxidized glutathione (1.2 mM). Kinetic data were fitted to a rectangular hyperbola and the kinetic constants evaluated as described previously¹⁷.

TABLE 3 Amino-acid comparisons of category I nicotinamide coenzyme-binding enzymes around the $\beta\alpha\beta$ dinucleotide-binding fold

NAD(PH) binding																			
Adrenodoxin reductase																			
Bovine	152	G	Q	G	N	V	A	L	D	V	A	R	I	L	L	T	P	P	
Human	151	G	Q	G	N	V	A	L	D	V	A	R	I	L	L	T	P	P	
Octopine synthase																			
<i>Agrobacterium</i>	8	G	A	G	N	V	A	L	T	L	A	G	D	L	A	R	R	L	G
Malic enzyme																			
Rat	300	G	A	G	E	A	A	L	G	I	A	H	L	I	V	M	A	M	
Glutamate dehydrogenase																			
<i>E. coli</i>	239	G	S	G	N	V	A	Q	Y	A	I	E	K	A	M	E	F	G	
<i>N. crassa</i>	225	G	S	G	N	V	A	Q	Y	A	A	L	K	L	I	E	L	G	
Yeast	224	G	S	G	N	V	A	Q	Y	A	A	L	K	V	I	E	L	G	
Mercuric reductase																			
<i>S. aureus</i> pl258	257	G	S	G	Y	I	A	A	E	L	G	Q	M	F	H	N	L	G	
<i>P. aeruginosa</i> Tn501	277	G	S	S	V	V	A	L	E	L	A	Q	A	F	A	R	L	G	
<i>S. flexnerii</i> R100	276	G	S	S	V	V	A	L	E	L	A	Q	A	F	A	R	L	G	
Trypanothione reductase																			
<i>T. congolense</i>	195	G	G	G	F	I	S	V	E	F	A	G	I	F	N	A	Y	K	P N G
Glutathione reductase																			
<i>E. coli</i>	174	G	A	G	Y	I	A	V	E	L	A	G	V	I	N	G	L	G	
Human	194	G	A	G	Y	I	A	V	E	M	A	G	I	L	S	A	L	G	
<i>P. aeruginosa</i>	172	G	G	G	Y	I	A	V	E	F	A	S	I	F	N	G	L	G	
Thioredoxin reductase																			
<i>E. coli</i>	152	G	G	G	N	T	A	V	E	E	A	L	Y	L	S	N	I	A	
NAD(H) binding																			
Dihydrolipoamide dehydrogenase																			
<i>E. coli</i>	180	G	G	G	I	L	G	L	E	M	G	T	V	Y	H	A	L	G	
<i>B. stearothermophilus</i>	183	G	G	G	Y	I	G	I	E	L	G	T	A	Y	A	N	F	G	
Yeast	209	G	G	G	I	I	G	L	E	M	G	S	V	Y	S	R	L	G	
Human	220	G	A	G	V	I	G	V	E	L	G	S	V	W	Q	R	L	G	
Pig	220	G	A	G	V	I	G	V	E	L	G	S	V	W	Q	R	L	G	
<i>A. vinelandii</i>	187	G	A	G	V	I	G	L	E	L	G	S	V	W	A	R	L	G	
<i>P. fluorescens</i>	188	G	A	G	V	I	G	L	E	L	G	S	V	W	A	R	L	G	
Alcohol dehydrogenase																			
<i>Drosophila</i>	14	G	L	G	G	I	G	L	D	T	S	K	Q	L	L	K	R	D	
Rat	15	G	L	G	G	V	G	L	S	V	V	I	G	C	K	T	A	G	
Lactate dehydrogenase																			
<i>B. stearothermophilus</i>	13	G	A	G	F	V	G	A	S	Y	V	F	A	L	M	N	Q	G	
<i>B. caldotenax</i>	13	G	R	G	F	V	G	A	S	Y	A	F	A	L	M	N	Q	G	
<i>Bacillus X₁</i>	12	G	A	G	F	V	G	S	S	Y	V	F	A	L	I	N	Q	G	
Mouse	25	G	V	G	D	V	G	M	A	C	A	I	S	I	L	L	K	G	
Pig, M-type	26	G	V	G	A	V	G	M	A	C	A	I	S	I	L	M	K	E	
Pig, H-type	27	G	V	G	Q	V	G	M	A	C	A	I	S	I	L	G	K	S	
Glyceraldehyde phosphate dehydrogenase																			
<i>B. stearothermophilus</i>	8	G	F	G	R	I	G	R	N	V	F	R	A	A	L	K	N	P	
<i>T. aquaticus</i>	7	G	F	G	R	I	G	R	Q	V	F	R	I	I	H	S	R	G	
Yeast	7	G	F	G	R	I	G	R	L	V	M	R	I	A	L	S	R	P	
Pig	7	G	F	G	R	I	G	R	L	V	T	R	A	A	F	N	S	G	

Residues shown in bold indicate the 'fingerprint' region identified by Wierenga *et al.*²⁵. The number indicates the position in the primary structure of the first residue shown for each sequence. The sequences listed are: adrenodoxin reductase (bovine and human); octopine synthase; malic enzyme; glutamate dehydrogenase (*E. coli*, *Neurospora crassa*, yeast); for individual references see ref. 26; mercuric reductase (*S. aureus*, ref. 34; *P. aeruginosa*, ref. 35; *S. flexnerii*, ref. 36); trypanothione reductase (*T. congolense*, ref. 37); glutathione reductase (*E. coli*, ref. 13; human, ref. 33; *P. aeruginosa*, N. L. Brown—personal communication); thioredoxin reductase (ref. 44); dihydrolipoamide dehydrogenase (*E. coli*, ref. 38; *B. stearothermophilus* (A. Borges, C. F. Hawkins & R. N. Perham—unpublished results); yeast, ref. 39; human and pig, ref. 40; *P. fluorescens*, ref. 41; *A. vinelandii*, ref. 42; alcohol dehydrogenase (*Drosophila* and rat); lactate dehydrogenase (*B. stearothermophilus*, *B. caldotenax*, *Bacillus X₁*, mouse, pig M-type and pig H-type); glyceraldehyde phosphate dehydrogenase (*B. stearothermophilus*, *Thermus aquaticus*, yeast and pig; for individual references, see ref. 25).

the enzyme with this single substitution and in enzymes already carrying one or both of the arginine mutations described above. This was not accompanied by great changes in k_{cat} (Table 2E). The A179G mutation on its own had little effect on the kinetic parameters of the enzyme for NADPH when combined with the mutations of either Arg 204 or Arg 204 and Arg 198 in the 2'-phosphate-binding site (Table 2B). By contrast, the K_m of the R198M mutant for NADPH did not decrease with the addition of the A179G mutation. The structural basis for this observation is at present unknown. The substitution of Ala 183 by glycine (A183G) was subsequently added to the A179G/R198M/R204L mutant and was found to have little further effect on the value of K_m or k_{cat} for NADH (Table 2E), although k_{cat} and K_m for NADPH both increased about fivefold

(Table 2B). Similarly, we introduced the mutation (R204P) of Arg 204 to proline, the residue found in the corresponding position in *E. coli* dihydrolipoamide dehydrogenase (Table 1), in place of the R204L mutation that had previously been used. The R204P mutation had little effect on the values of K_m and k_{cat} for NADPH or NADH in the enzyme already carrying the three mutations A179G, A183G and R198M (Table 2B and E).

Substitution of Val 197, Lys 199 and His 200. In *E. coli* dihydrolipoamide dehydrogenase, the negatively charged residue at the end of the second β -strand of the $\beta\alpha\beta$ -fold which forms a hydrogen bond to the 2'-hydroxyl group of the adenine ribose of NAD⁺ (refs 25 and 31) is Glu 203 (Table 1). The corresponding residue in *E. coli* glutathione reductase is Val 197; we there-

fore investigated the effect of replacing this residue with glutamic acid (V197E). To remove any residual positive charge in the 2'-phosphate-binding site of NADP⁺ in *E. coli* glutathione reductase and to mimic yet more closely the NAD⁺-binding site of *E. coli* dihydrolipoamide dehydrogenase (Table 1), we created two further mutations: Lys 199 to phenylalanine (K199F) and His 200 to aspartic acid (H200D). All three mutations were simultaneously introduced into the *E. coli gor* gene already carrying the A179G, A183G, R198M and R204P mutations. These three additional replacements raised the K_m for NADPH by about threefold, left the K_m for NADH almost unchanged, and raised the k_{cat} for NADH by about threefold (Table 2C and F). This protein, with seven differences in amino-acid sequence compared with wild-type glutathione reductase, showed a marked preference for NADH rather than NADPH as coenzyme.

Discussion

Perhaps the most convenient measure of the discrimination of an enzyme for its substrate is the ratio of k_{cat} to K_m (ref. 32).

The effects of the mutations described here on this ratio are displayed in Fig. 3. The value of the k_{cat} to K_m ratio for NADPH is lowered by a factor of ~250, whereas that for NADH is increased by a factor of ~70. The ratio of these two contrasting shifts is 18,000—evidence of the profound switch in coenzyme specificity that has been achieved. Given that the molecular basis of binding is much better understood than that of catalysis, it is gratifying that a main component of the switch is embodied in changes in k_{cat} .

It is important to note that the mutations we have identified as crucial to the change of coenzyme specificity are all located in the $\beta\alpha\beta$ -fold fingerprint region^{25,26} of the nicotinamide nucleotide-binding domain of glutathione reductase. So it is probable that similar experiments can now be performed to redesign the coenzyme specificity of any nicotinamide nucleotide-specific enzyme that contains an homologous dinucleotide-binding domain. To tune more finely the discrimination in favour of NADH will require better knowledge of the stereochemical effects of the mutations, probably obtainable only by X-ray crystallographic analysis. □

Received 13 September; accepted 24 November 1989.

1. Stryer, L. *Biochemistry* (Freeman, New York, 1988).
2. Graf, L. *et al. Biochemistry* **26**, 2616–2623 (1987).
3. Cronin, C. N., Malcolm, B. A. & Kirsch, J. F. *J. Am. chem. Soc.* **109**, 2222–2223 (1987).
4. Oxender, D. L. & Fox, C. F. *Protein Engineering* (Liss, New York, 1987).
5. Lindberg, R. L. P. & Negishi, M. *Nature* **339**, 632–634 (1989).
6. Clarke, A. R., Atkinson, T. & Holbrook, J. J. *Trends Biochem. Sci.* **14**, 101–105, 145–148 (1989).
7. Holmgren, A. *Rev. Biochem.* **54**, 237–271 (1985).
8. Schirmer, R. H., Krauth-Siegel, R. L. & Schulz, G. E. in *Glutathione: Chemical, Biochemical and Medical Aspects. Part A* (eds Dolphin, D., Poulson, R. & Avramovic, O.) 553–596 (Wiley, New York, 1989).
9. Chung, Y. C. & Hurlbert, R. E. *J. Bact.* **123**, 203–211 (1975).
10. Thieme, R., Pai, E. F., Schirmer, R. H. & Schulz, G. E. *J. molec. Biol.* **151**, 763–782 (1981).
11. Pai, E. F. & Schulz, G. E. *J. molec. Biol.* **258**, 1751–1757 (1983).
12. Karplus, P. A. & Schulz, G. E. *J. molec. Biol.* **195**, 701–729 (1987).
13. Greer, S. & Perham, R. N. *Biochemistry* **25**, 2736–2742 (1986).
14. Scrutton, N. S., Berry, A. & Perham, R. N. *Biochem. J.* **245**, 875–880 (1987).
15. Deonarain, M. P., Berry, A., Scrutton, N. S. & Perham, R. N. *Biochemistry* (in the press).
16. Scrutton, N. S., Berry, A. & Perham, R. N. *FEBS Lett.* **241**, 46–50 (1988).
17. Berry, A., Scrutton, N. S. & Perham, R. N. *Biochemistry* **28**, 1264–1269 (1989).
18. Perham, R. N., Harrison, R. A. & Brown, J. P. *Biochem. Soc. Trans.* **6**, 47–50 (1978).
19. Carothers, D. J., Pons, G. & Patel, M. S. *Archs Biochem. Biophys.* **268**, 409–425 (1989).
20. Takenaka, A. *et al. J. Biochem.* **103**, 463–469 (1988).
21. Schierbeek, A. J. *et al. J. molec. Biol.* **206**, 365–380 (1989).
22. Rice, D. W., Schulz, G. E. & Guest, J. R. *J. molec. Biol.* **174**, 483–496 (1984).
23. Pai, E. F., Karplus, P. A. & Schulz, G. E. *Biochemistry* **27**, 4465–4474 (1988).
24. Rossmann, M. G., Liljas, A., Brändén, C. I. & Banaszak, L. J. *Enzymes* **11**, 61–102 (1975).
25. Wierenga, R. K., de Maeyer, M. C. H. & Hol, W. G. J. *Biochemistry* **24**, 1346–1357 (1985).
26. Hanukoglu, I. & Gutfinger, T. *Eur. J. Biochem.* **180**, 479–484 (1989).
27. Feeney, J., Birdsall, B., Roberts, G. C. K. & Burgen, A. S. V. *Nature* **257**, 564–566 (1975).
28. Creighton, T. E. *Proteins: Structures and Molecular Principles* (Freeman, New York, 1984).
29. Brändén, C. I. *Q. Rev. Biophys.* **13**, 317–338 (1980).
30. Hol, W. G. J., Van Duijn, P. Th. & Berendsen, H. J. C. *Nature* **273**, 443–446 (1978).
31. Adams, M. J. *et al.* in *Structural Studies on Molecules of Biological Interest* (eds Dodson, G., Glusker, J. P. & Sayre, D.) 328–338 (Oxford University Press, 1981).
32. Fersht, A. R. *Enzyme Structure and Mechanism* 2nd edn (Freeman, New York, 1985).
33. Krauth-Siegel, R. L. *et al. Eur. J. Biochem.* **121**, 259–267 (1982).
34. Laddaga, R. A., Chu, L., Misra, T. K. & Silver, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5106–5110 (1987).
35. Brown, N. L., Ford, S. J., Pridmore, D. & Fritzinger, D. C. *Biochemistry* **22**, 4089–4095 (1983).
36. Misra, T. K. *et al. Gene* **34**, 253–262 (1985).
37. Shames, S. L., Kimmel, B. E., Peoples, O. P., Agabian, N. & Walsh, C. T. *Biochemistry* **27**, 5014–5019 (1988).
38. Stephens, P. E., Lewis, H. M., Darlison, M. G. & Guest, J. R. *Eur. J. Biochem.* **135**, 519–527 (1983).
39. Browning, K. S., Uhlinger, D. J. & Reed, L. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1831–1834 (1988).
40. Otulakowski, K. G. & Robinson, B. H. *J. biol. Chem.* **262**, 17313–17318 (1987).
41. Benen, J. A. E., van Berkel, W. J. H., van Dongen, W. M. A. M., Müller, F. & de Kok, A. *J. gen. Microbiol.* **135**, 1787–1797 (1989).
42. Westphal, A. H. & de Kok, A. *Eur. J. Biochem.* **172**, 299–305 (1988).
43. Taylor, J. W., Schmidt, W., Kosstick, R., Okruszek, A. & Eckstein, F. *Nucleic Acids Res.* **13**, 8749–8764 (1985).
44. Russel, M. & Model, P. *J. biol. Chem.* **263**, 9015–9019 (1988).

ACKNOWLEDGEMENTS. We thank Nigel W. Hinchliffe for skilled technical assistance. This work was supported by the SERC and its Cambridge Centre for Molecular Recognition, the Royal Society and the Royal Commission for the Exhibition of 1851. A.B. is a Royal Society 1983 University Research Fellow. N.S.S. is a Research Fellow of the Royal Commission for the Exhibition of 1851.

LETTERS TO NATURE

Spectral differences between radio galaxies and quasars

N. Jackson & I. W. A. Browne

Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

It has been suggested recently^{1–4} that two classes of radio-loud extragalactic object, narrow-line radio galaxies and quasars, are intrinsically similar, differing only in the angle from which they are seen by the observer. In the simplest version of such models, one would expect that the [O III] 5,007-Å line luminosities of the two types of object would be similar for a given extended radio luminosity and redshift. In spectroscopic observations of a sample of radio galaxies and quasars selected so as to be otherwise comparable, we find that this is not the case: the quasar [O III] luminosities exceed those of radio galaxies by factors of 5–10. Extinction by at least two magnitudes as a result of orientation differences is necessary for these observations to remain consistent with the radio galaxy/quasar unification hypothesis; we suggest

that evidence for such extinction may exist in the form of high H α /H β ratios in radio galaxies^{5–8}.

Steep-spectrum (radio-lobe-dominated) quasars and high-luminosity radio galaxies are two classes of extragalactic object with radio structures that are often strikingly similar to one another. There are many examples of both types of objects, which consist of a compact core surrounded by two diffuse radio lobes on either side of the core. Core-dominated quasars (CDQs) do not conform to this pattern: they consist of a bright core surrounded by a fainter nebulosity. It has been argued^{9–11} that this difference is caused by orientation effects: specifically, that CDQs are lobe-dominated quasars (LDQs) seen with their radio axes (that is, their radio jets) pointing close to the line of sight, so that the core flux is amplified by Doppler boosting. The unbeamed (or only slightly beamed¹²) lobe emission then appears as a fainter halo or extension from the bright core. If this model is correct, extended radio luminosity is an indication of the intrinsic power of the source.

It has been proposed^{1,3} that powerful (Fanaroff–Riley¹³ class 2) narrow-line radio galaxies can be brought into this scheme, if their axes lie even closer to the sky plane ($\leq 45^\circ$; ref. 3) than those of LDQs. But quasars have a strong optical continuum

TABLE 1 Observed parameters for our sample

Radio galaxy	z	log L_{ext}	log $L[\text{O III}]$	Quasar	z	log L_{ext}	log $L[\text{O III}]$
3C28	0.195	25.9 (refs 23, 24)	<33.75	2135-147	0.200	26.0 (ref. 25)	36.38
3C33.1	0.181	26.1 (refs 24, 26)	35.05	3C206	0.200	26.1 (ref. 27)	35.76
3C42	0.395	26.9 (refs 26, 28)	35.29	3C351	0.371	26.9 (ref. 25)	36.71 (ref. 29)
3C46	0.44	26.6 (refs 24, 26)	36.04	3C215	0.411	26.6 (ref. 25)	35.82
3C173.1	0.29	26.5 (refs 26, 28)	34.72 (ref. 30)	3C249.1	0.311	26.6 (ref. 25)	36.64
3C220.1	0.61	27.1 (refs 26, 28)	36.00	1007+417	0.613	27.1 (ref. 31)	36.78
3C265	0.81	27.5 (refs 26, 28)	37.05	3C175	0.768	27.5 (ref. 25)	36.85
3C274.1	0.42	26.9 (refs 28, 32)	34.59	3C47	0.425	26.8 (ref. 25)	36.45
3C284	0.239	26.3 (refs 23, 24)	35.59	1004+130	0.240	26.2 (ref. 25)	35.95
3C300	0.270	26.6 (refs 28, 32)	35.57	2201+315	0.298	26.6 (ref. 25)	36.02
3C436	0.22	26.4 (refs 23, 28)	35.09 (ref. 30)	3C323.1	0.264	26.5 (ref. 25)	36.27 (ref. 29)
3C438	0.29	26.8 (refs 26, 28)	34.47 (ref. 30)	3C277.1	0.320	26.8 (ref. 25)	36.41

Extended radio luminosity L_{ext} is given in W Hz^{-1} at 5 GHz, [O III] luminosities are given in watts ($H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $q_0 = 0$). [O III] luminosities without citations are our own observations. 3C28 is listed¹⁵ as a Fanaroff-Riley class 1 (low-power) source¹³, but because of its uncertain radio morphology^{33,34} and high radio luminosity³³ compared to many Fanaroff-Riley class 2 (high-power) sources we have included it here.

and broad emission lines whereas narrow-line radio galaxies do not. (We shall use radio galaxy to refer to narrow-line radio galaxies only; broad-line radio galaxies do not concern us here, because they have many of the properties of low-luminosity, nearby quasars.) To solve this problem we invoke a region of obscuring material distributed around the core of the radio galaxy in such a way as to hide the continuum and broad lines emitted from the object's central region when the object lies close to the sky plane. In a simple model it would not be expected that this obscuration would affect the narrow lines, which are emitted from a region that is a considerable distance from the centre of the object. One would, however, expect that quasars should have a higher optical non-thermal luminosity for a given [O III] luminosity than do radio galaxies. It has been shown¹⁴ that this is the case, although both quasars and radio galaxies show separate correlations of [O III] luminosity with non-thermal luminosity. Therefore the analysis of optical data is consistent with the unification model. Here we try to establish whether those quasars and radio galaxies that the scheme would predict to be intrinsically similar actually have indistinguishable narrow-line luminosities.

To make a comparison between the narrow lines of quasars and radio galaxies, we have obtained spectra of quasars and radio galaxies with the William Herschel and Isaac Newton telescopes on La Palma. The radio galaxies have been chosen from the 3CRR catalogue¹⁵ to have redshifts less than 0.85, to permit observation of the [O III] line. For each radio galaxy a quasar has been chosen to match it in extended radio luminosity

to within a factor of 30% and in redshift to within 0.05. This process gives a possible 18 pairs in which each quasar and radio galaxy is used only once: because of adverse observing conditions we have so far been able to observe only nine of them. By matching in radio luminosity we compare objects of the same intrinsic power (see above). This is important because it is known that for radio galaxies¹⁶, and possibly for quasars¹⁷, there is a correlation between narrow-line luminosity and extended radio emission. Early comparisons of the emission-line properties of radio galaxies and quasars did not take this radio-power dependence into account and are therefore not directly relevant here. By redshift matching we exclude effects of any cosmological evolution from the comparison. We have used similar slit widths (2–2.4 arcsec) for both the radio galaxies and the quasars.

The results are summarized in Table 1. It is clear that, as previously suspected¹⁸, the [O III] luminosities of the two classes of object are strikingly different. The quasar [O III] lines are typically stronger by a factor of about ten than those of radio galaxies. In the nine pairs of objects for which we have observed the radio galaxies, a Wilcoxon signed rank test allows the rejection at the 98% level of the hypothesis that the two distributions are drawn from the same parent population (Fig. 1). If three further pairs found by searching the literature are included (Table 1) the hypothesis that the two types of object have the same [O III] luminosity can be rejected at a level of $\gg 99\%$. Therefore it seems that, with few exceptions (3C265, for example), radio galaxies have weaker [O III] lines than do quasars of the same redshift and extended radio power.

Can we reconcile the radio galaxy/quasar unified scheme with these data? To do so we have to postulate orientation-dependent dust extinction of 2–3 magnitudes within the narrow-line region, which should affect the flux observed along our line of sight only in radio galaxies. We should then expect independent evidence for extinction in radio galaxies, and such evidence exists. The Balmer-line ratios in narrow-line radio galaxies have been studied extensively, and the ratios of $H\alpha$ to $H\beta$ are generally found^{5–8} to be greater than the simple recombination expectations¹⁹ (although larger theoretical values are possible²⁰). It is generally believed that these Balmer-line ratios and other indicators are consistent with dust extinction of two magnitudes at 5,000 Å in many radio galaxies. Also, there should not be an equal amount of extinction in the narrow-line regions of quasars. The observational evidence available is not good enough to address this question quantitatively, though predominantly blue asymmetries observed in the [O III] lines of core-dominated quasars by others^{21,22} and by us (unpublished observations) may indicate that some dust is present. (Asymmetries can be produced by a combination of radial outflow of emitting material and extinction). A careful programme to look for reddening in the narrow-line regions of quasars is required. □

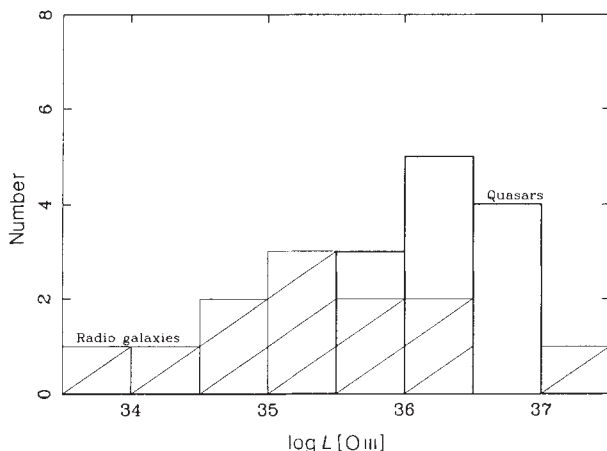


FIG. 1 Histogram of [O III] luminosities in high-power radio galaxies and quasars.

Received 29 August; accepted 13 November 1989.

1. Scheuer, P. A. G. in *Superluminal Radio Sources* (eds Zensus, J. A. & Pearson, T. J.) 104–113 (Cambridge University Press, 1987).
2. Morisawa, K. & Takahara, F. *Mon. Not. R. astr. Soc.* **228**, 745–758 (1987).
3. Barthel, P. D. *Astrophys. J.* **336**, 606–611 (1989).
4. Browne, I. W. A. in *BL Lac Objects – 10 Years After* (eds Maraschi, L., Maccacaro, T. & Ulrich, M.-H.) 401–411 (Springer, Berlin, 1989).
5. Costero, R. & Osterbrock, D. E. *Astrophys. J.* **211**, 675–683 (1977).
6. Koski, A. T. *Astrophys. J.* **223**, 56–73 (1978).
7. Gaskell, C. M. *Publ. Astr. Soc. Pacif.* **94**, 891–893 (1984).
8. Ferland, G. J. & Osterbrock, D. E. *Astrophys. J.* **289**, 105–108 (1985).
9. Blandford, R. D. & Königl, A. *Astrophys. J.* **232**, 34–48 (1979).
10. Orr, M. J. L. & Browne, I. W. A. *Mon. Not. R. astr. Soc.* **200**, 1067–1082 (1982).
11. Jackson, N., Browne, I. W. A., Murphy, D. W. & Saikia, D. J. *Nature* **338**, 485–487 (1989).
12. Murphy, D. W. thesis, Manchester Univ. (1988).
13. Fanaroff, B. L. & Riley, J. M. *Mon. Not. R. astr. Soc.* **167**, 31P–35P (1974).
14. Yee, H. K. C. *Astrophys. J.* **241**, 894–902 (1980).
15. Laing, R. A., Riley, J. M. & Longair, M. S. *Mon. Not. R. astr. Soc.* **204**, 151–187 (1983).
16. Saunders, R., Baldwin, J. E., Rawlings, S., Warner, P. J. & Miller, L. *Mon. Not. R. astr. Soc.* **238**, 777–790 (1989).
17. Willis, B. & Browne, I. W. A. *Astrophys. J.* **302**, 56–63 (1986).
18. Baum, S. A. & Heckman, T. M. *Astrophys. J.* **336**, 702–721 (1989).
19. Osterbrock, D. E. *Astrophysics of Gaseous Nebulae* (Freeman, San Francisco, 1974).
20. Avrett, E. H. & Loesser, R. *Astrophys. J.* **331**, 211–246 (1988).
21. Heckman, T. M., Miley, G. K. & Green, R. F. *Astrophys. J.* **281**, 525–534 (1984).
22. de Waard, G. thesis, Univ. Leiden (1986).
23. Riley, J. M. & Pooley, G. G. *Mem. R. astr. Soc.* **80**, 105–137 (1975).
24. Kellermann, K. I., Pauliny-Toth, I. I. K. & Williams, P. J. S. *Astrophys. J.* **157**, 1–34 (1969).
25. Browne, I. W. A. & Murphy, D. *Mon. Not. R. astr. Soc.* **226**, 601–627 (1987).
26. Jenkins, C. J., Pooley, G. G. & Riley, J. M. *Mem. R. astr. Soc.* **84**, 61–99 (1977).
27. Miley, G. K. & Hartsuiker, A. P. *Astr. Astrophys. Suppl.* **34**, 129–163 (1978).
28. Kühn, H., Nauber, U., Pauliny-Toth, I. I. K. & Witzel, A. Max-Planck-Institut für Radioastronomie preprint no. 55 (1979).
29. Boroson, T. A. & Oke, J. B. *Astrophys. J.* **281**, 535–544 (1984).
30. Rawlings, S. thesis, Univ. Cambridge (1987).
31. Owen, F. N. & Puschell, J. J. *Astr. J.* **89**, 932–957 (1984).
32. Leahy, J. P. & Williams, A. G. *Mon. Not. R. astr. Soc.* **210**, 929–951 (1984).
33. Macklin, J. T. *Mon. Not. R. astr. Soc.* **203**, 147–155 (1983).
34. Feretti, L., Gioia, I. M., Giovannini, G., Gregorini, L. & Padrieli, L. *Ast. Astrophys.* **139**, 50–54 (1984).

ACKNOWLEDGEMENTS. The Isaac Newton and William Herschel telescopes are operated on the island of La Palma by the Royal Greenwich Observatory at the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. We thank P. Charles and J. Cepa for support at the telescopes, R. Watson for software assistance and D. Murphy for comments.

A supercluster of IRAS galaxies behind the Great Attractor

David M. Allen*, R. P. Norris†, L. Staveley-Smith*, V. S. Meadows* & P. F. Roche*‡

* Anglo-Australian Observatory, PO Box 296, Epping, New South Wales 2121, Australia

† Australia Telescope National Facility, PO Box 76, Epping, New South Wales 2121, Australia

THE dynamics of nearby galaxies support the existence of a large concentration of mass (the ‘Great Attractor’) in the Centaurus region at a redshift (z) between 0.012 and 0.015 (refs 1, 2). Scaramella *et al.*³ have suggested that a concentration of rich clusters of galaxies at $z \approx 0.05$ in the Centaurus region may also have a significant role in perturbing the local Hubble flow. The existence of such overdense regions in the Universe would have a profound effect on ideas about the formation of large-scale structure⁴. At present the reality of such regions is supported by some, but not by all, studies of the streaming of distant galaxies^{1,2,5–9}. We demonstrate here that redshifts of galaxies that are in the Infrared Astronomical Satellite (IRAS) survey show the existence of an extended supercluster at $z \approx 0.05$ surrounding the concentration of clusters described by Scaramella *et al.* Its contribution to the Local Group peculiar (non-Hubble) velocity is probably minor, although it may dominate motions on the far side of the Great Attractor.

IRAS galaxies are regarded as good tracers of mass in the universe, especially because their surface-density distribution on the sky can be fitted by a dipole the axis of which coincides with that of the cosmic microwave background^{10–13} when viewed from the Local Group barycentre. We are using the Anglo-

Australian Telescope in a redshift survey, which we will report elsewhere, of IRAS galaxies in three separate fields. In each field we have defined a sample of infrared-luminous galaxies that have been detected by IRAS¹⁴ at both 60 and 100 μm . Using the redshifts we also select galaxies with infrared luminosities L_{IR} (calculated as in ref. 15 using $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and a bolometric correction of 1.45) in excess of $4 \times 10^{10} L_{\odot}$, and so exclude normal spiral galaxies. The IRAS sensitivity ensures completeness at this luminosity to $z \sim 0.06$.

The concentration of galaxies at $z \sim 0.05$ in Centaurus was first noticed 60 years ago¹⁶, but is only now being quantified¹⁷. We shall refer to it as the Centaurus distant supercluster (CDS) to distinguish it from the more local Centaurus-Pavo (C-P) supercluster widely associated with the Great Attractor. Redshifts are not available for the majority of clusters in the CDS with $z > 0.02$, so Scaramella *et al.*³ used the diameter of the tenth brightest galaxy in each to estimate distance. In the region of space bounded by $12^{\text{h}}40^{\text{m}} < \text{RA} < 14^{\text{h}}08^{\text{m}}$, $-41^{\circ} < \text{dec} < -25^{\circ}$, (RA, right ascension; dec, declination) and $z < 0.067$ they found an overdensity of rich clusters of a factor of about ten. They dubbed this volume the α -region. Within this volume, the majority of clusters were estimated to lie in the redshift range $0.03 < z < 0.06$.

Our field 1 ($12^{\text{h}}30^{\text{m}} < \text{RA} < 14^{\text{h}}30^{\text{m}}$, $-40^{\circ} < \text{dec} < -17^{\circ}$) encompasses virtually the entire α -region and within the overlapping region of right ascension and declination we find 29 galaxies with $L_{\text{IR}} > 4 \times 10^{10} L_{\odot}$ and $0.03 < z < 0.06$. Our redshift survey is 70% complete for all 413 IRAS galaxies that we have identified in field 1, and $> 95\%$ complete for galaxies with diameters < 1 arcmin. We examined the redshift-diameter relationship for IRAS galaxies in this field and used it to estimate that about four further galaxies for which the redshifts have not been determined may populate the region $0.05 < z < 0.06$. Our present analysis is not significantly affected by this slight incompleteness. We have also obtained redshifts to comparable depth and completeness in a field 3a separated by exactly 12 hours from field 1 and covering the same declination range.

Figure 1 presents histograms of the redshift distribution of a number of samples. In Fig. 1a we show the distribution of all

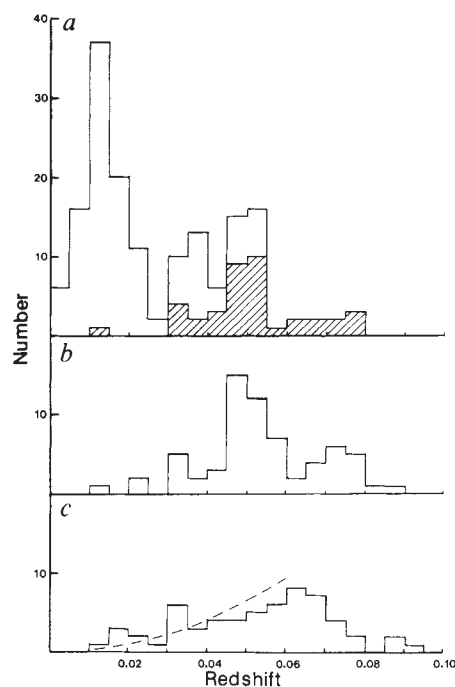


FIG. 1 Redshift histograms. a, Galaxies in the α -region. Open boxes: all galaxies with redshifts; hatched: galaxies with $L_{\text{IR}} > 4 \times 10^{10} L_{\odot}$. b, Galaxies with $L_{\text{IR}} > 4 \times 10^{10} L_{\odot}$ in the entirety of our field 1. c, As in b for our field 3a, showing that the distribution of galaxies matches well a uniform-density model, indicated by the broken line.

‡ Present address: Department of Astrophysics, Oxford University, Keble Road, Oxford OX1 3RH, UK.

IRAS galaxies measured in the α -region; those associated with the C-P supercluster are apparent at $z \sim 0.013$. We also show (hatched in Fig. 1a) the subset of 39 galaxies with $L_{\text{IR}} > 4 \times 10^{10} L_{\odot}$ and $z < 0.10$. A peak is seen within the range $0.045 < z < 0.055$, whereas the foreground supercluster is no longer a recognizable feature. Figure 1b shows the redshift histogram for galaxies with $L_{\text{IR}} > 4 \times 10^{10} L_{\odot}$ in the entirety of our field 1 and, for comparison, Fig. 1c shows the corresponding histogram for field 3a. The peak at $0.045 < z < 0.055$ is even more apparent in the entirety of field 1 than in the α -region alone and, indeed, extends to slightly higher redshift. By contrast, the distribution in field 3a is well fitted by a uniform space density of galaxies to $z = 0.06$, beyond which it falls steeply because of incompleteness in the IRAS survey. A Kolmogorov-Smirnov test reveals that fields 1 and 3a differ in their redshift distributions at the 97% confidence level. A detailed comparison of the two fields is not straightforward, because IRAS scanned field 3a twice whereas field 1 was covered three times, and the confusion that is due to galactic cirrus and crowding is worse in field 1. We estimate an overdensity of IRAS galaxies with the redshift range $0.045 < z < 0.055$ in field 1 of $\delta\rho/\rho_0 \sim 2$, but regard the number as provisional pending a more comprehensive survey of redshifts. Uncertainty in the bright end of the IRAS luminosity function also precludes a more quantitative analysis.

The IRAS galaxy sample shows other regions of clumping; one may also lie in our field 1 at $0.073 < z < 0.08$, $\text{RA} > 13^{\text{h}}40^{\text{m}}$, as Figs 1b and 2 show. The IRAS data are complete to this redshift for $L_{\text{IR}} > 8 \times 10^{10} L_{\odot}$. For such galaxies the distribution in field 1 differs from a uniform model at the 90% confidence level at $z = 0.073$. The reality of this second group may be confirmed when our redshift survey is complete.

The detailed distribution of IRAS galaxies within the α -region shows that most do not lie within the bounds of the rich clusters. This is to be expected because most IRAS galaxies are gas rich, and rich clusters are deficient in gas-rich galaxies. The excess of luminous IRAS galaxies in the α -region is also seen on Fig. 2, which shows pie diagrams in right ascension and declination for all galaxies with $L_{\text{IR}} > 4 \times 10^{10} L_{\odot}$ within our field 1. A similar pie diagram is shown for field 3a. The concentration of IRAS

galaxies in the relevant redshift range extends beyond the α -region in declination, so that the clusters are only the central manifestation of a more extended mass concentration that is due to the surrounding CDS. This feature spans at least 60 Mpc, making it one of the largest known superclusters¹⁸.

Judging from the luminous IRAS galaxies alone, the mass concentration in the CDS is much greater than that in the more local C-P supercluster. Its greater distance, however, diminishes its effect on the peculiar velocity of the Local Group. If our estimates of size and overdensity represent the true distribution of matter in the region, then the resultant peculiar velocity of the Local Group¹⁹ is $\sim 80\Omega_0^{0.6} \text{ km s}^{-1}$ for a cosmological density parameter Ω_0 . This is a small, but not negligible, component of the total Local Group velocity and could partly explain why infall towards the C-P supercluster seems to dominate² over that towards the apparently equally massive Perseus-Pisces supercluster which lies in the opposite direction. The corollary of this is that in the region between the C-P supercluster and the CDS, there will be a cancellation of gravitational attraction. We estimate that infall towards the CDS will dominate that into the far side of the C-P supercluster for $z \geq 0.025$ if $\Omega_0 = 1$. If continuing studies to measure streaming motions on the far side of the Great Attractor confirm the large mass concentration in the CDS, then cold-dark-matter⁴ theories of galaxy formation will be in extreme difficulty. Large bulk flows over a scale $\Delta z \approx 0.05$ are very improbable in flat or open Universes dominated by such matter²⁰. $\square$

Received 19 July; accepted 13 November 1989.

- Lynden-Bell, D. *et al. Astrophys. J.* **326**, 19–49 (1988).
- Staveley-Smith, L. & Davies, R. D. *Mon. Not. R. astr. Soc.* **241** (in the press).
- Scaramella, R., Baiesi-Pillastrini, G., Chincarini, G., Vettolani, G. & Zamorani, G. *Nature* **338**, 562–564 (1989).
- Bertschinger, E. & Juszkiewicz, R. *Astrophys. J.* **334**, L59–L62 (1988).
- Rubin, V. C., Thonnard, N., Ford, W. K. & Roberts, M. S. *Astr. J.* **81**, 719–737 (1976).
- Peterson, C. J. & Baumgart, C. W. *Astr. J.* **91**, 530–538 (1986).
- Collins, C. A., Joseph, R. D. & Robertson, N. A. *Nature* **320**, 506–508 (1986).
- Aaronson, M. *et al. Astrophys. J.* **302**, 536–563 (1986).
- Lucy, J. R. & Carter, D. *Mon. Not. R. astr. Soc.* **231**, 15P–20P (1988); **235**, 1177–1201 (1988).
- Yahil, A., Walker, D. & Rowan-Robinson, M. *Astrophys. J.* **301**, L1–L5 (1986).
- Meiksin, A. & Davis, M. *Astr. J.* **91**, 191–198 (1986).
- Harmon, R. T., Lahav, O. & Meurs, E. J. A. *Mon. Not. R. astr. Soc.* **228**, 5P–10P (1987).
- Lahav, O., Rowan-Robinson, M. & Lynden-Bell, D. *Mon. Not. R. astr. Soc.* **234**, 677–701 (1988).
- The Joint IRAS Science Working Group *The Point Source Catalog* (U.S. Government Printing Office, Washington 1988).
- Lonsdale, C. J., Helou, G., Good, J. C. & Rice, W. *Catalogued Galaxies and Quasars Observed in the IRAS Survey* (Jet Propulsion Laboratory, Pasadena, 1985).
- Shapley, H. *Harvard Obs. Bull.* **874**, 9–12 (1930).
- Melnick, J. & Moles, M. *Rev. Mex. Astr. Astrofis.* **14**, 72–76 (1987).
- Bahcall, N. A. & Soneira, R. M. *Astrophys. J.* **277**, 27–37 (1984).
- Villumsen, J. V. & Davis, M. *Astrophys. J.* **308**, 499–509 (1986).
- Vittorio, N., Juszkiewicz, F. & Davis, M. *Nature* **323**, 132–133 (1986).

ACKNOWLEDGEMENTS. We thank R. Stathakis for her help in reducing some of the redshift data.

Determination of valence and cation distributions by resonant powder X-ray diffraction

J. P. Attfield

Chemical Crystallography Laboratory, University of Oxford,
9 Parks Road, Oxford OX1 3PD, UK

X-RAY crystallographic methods are important in determining the ordered atomic structures of matter from single-crystal or powder¹ diffraction data. But because the scattering factor depends on atomic number, it is not always possible to distinguish between atoms with very similar numbers of electrons, such as neighbouring elements or different valence states of the same element. By selecting an X-ray wavelength at an elemental absorption edge, the scattering power of that element is drastically altered by anomalous dispersion. Here I show that the resulting resonant

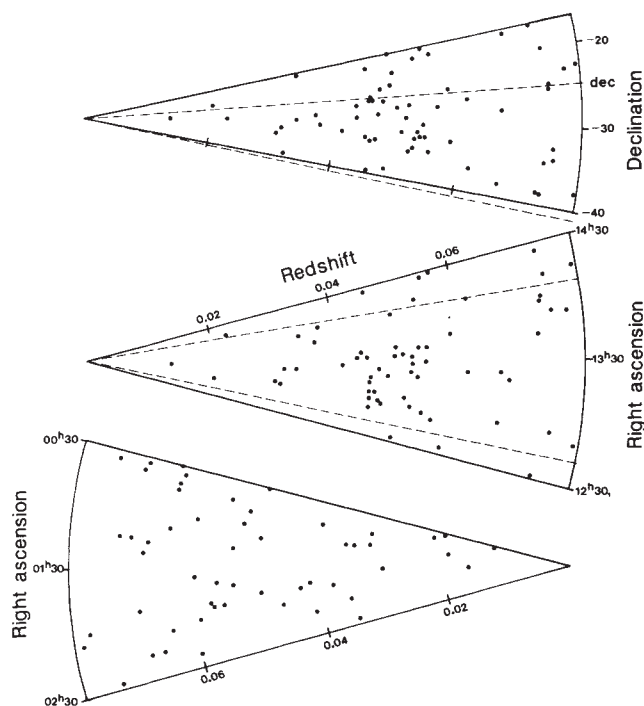


FIG. 2 Pie diagrams of luminous ($L_{\text{IR}} > 4 \times 10^{10} L_{\odot}$) IRAS galaxies in our fields 1 and 3a. The former is shown in both RA and dec, and the bounds of the α -region of ref. 3 are marked by broken lines.

X-ray diffraction data can provide information about valence states and cation distributions. I use powder data collected at the europium L_{III} edge using the SERC Daresbury Synchrotron Radiation Source to study the valence distribution in Eu_3O_4 and to determine the composition and cation and valence ordering in a new compound, EuSm_2O_4 . Such resonant X-ray diffraction methods should be applicable to almost all elements heavier than calcium for contrasting different electronic states of the resonant element and for differentiating it from neighbouring elements.

An atomic X-ray scattering factor f may be written as $f = f^n + f' + if''$, where f^n is the normal term which depends on the electron distribution, and f' and f'' are the real and imaginary parts of the anomalous scattering. The latter terms change rapidly with wavelength around an absorption edge and may reach very large magnitudes—values such as $f' = -32$ and $f'' = 31$ electrons per atom have been reported² at the L_{III} edges of Gd^{3+} and Sm^{3+} . Anomalous scattering effects are routinely used to elucidate complex non-centrosymmetric structures³ and to determine absolute configurations⁴ from single-crystal X-ray data, but had not been applied in powder diffraction work until the advent of variable-wavelength synchrotron powder X-ray diffractometers. Such instruments have recently been used to determine $f''(\text{Yb})$ in polycrystalline Yb_2O_3 close to the L_{III} absorption edge⁵, and to separate the temperature factors for Y^{3+} and Zr^{4+} sharing the same site in $\text{Zr}_{0.81}\text{Y}_{0.19}\text{O}_{1.90}$ from the anomalous scattering at three wavelengths around the metal K edges⁶.

As the L_{III} edges give rise to large anomalous scattering effects and occur at convenient wavelengths for the lanthanide metals, I have explored the possibilities of characterizing polycrystalline inorganic solids using these effects in a diffraction experiment on two Eu_3O_4 -type lanthanide oxides. Such materials have many interesting electronic and catalytic^{7,8} properties. Geometric arguments and cation substitution experiments have previously shown that Eu_3O_4 is a valence-ordered compound with three crystallographically distinct cation sites⁹. Eu^{3+} occupies the Eu(1) and Eu(2) sites and Eu^{2+} fills the Eu(3) position.

I prepared the samples using europium metal as a reducing agent. Mixtures of 1 Eu:4 Eu_2O_3 and 1 Eu:1 Eu_2O_3 :3 Sm_2O_3 were sealed in evacuated silica-glass tubes and heated at 900 °C for two days. The first reaction gave a mixture of orthorhombic Eu_3O_4 (ref. 10) and cubic C- Eu_2O_3 (ref. 11) and the second product contained a new Eu_3O_4 -type phase and monoclinic B- Sm_2O_3 (ref. 12). The presence of unreacted Ln_2O_3 (where Ln denotes a general lanthanide) in both samples resulted from the loss of some europium metal through reaction with the silica tube.

Synchrotron X-ray diffraction data were collected using the Synchrotron Radiation Source (SRS) instrument 8.3. A channel-cut Si (111) double-crystal monochromator was used to select a wavelength of 1.7771(1) Å (6,976.9 eV), calibrated by the positions of the C- Eu_2O_3 peaks ($a = 10.8683(2)$ Å, ref. 11) in the first sample. The count rate for the $\theta/2\theta$ scans of a rotating flat-plate sample was 2 s per 0.01° step in the ranges 20.00–97.23° for the europium oxides sample and 20.00–123.36° for the europium samarium oxides sample. The data were normalized by the incident-beam monitor count and scaled to counts per s per 100-mA ring current. The data were fitted using a standard multi-phase Rietveld program^{13,14} in which peak shapes are described by a 'pseudo-Voigt' function. I chose a fixed, overall, isotropic temperature factor of 0.10 Å² throughout to avoid excessive correlations between the refined parameters.

The first pattern was fitted by simultaneously refining the structures of Eu_2O_3 (ref. 15) and Eu_3O_4 (ref. 9) in space groups $Ia\bar{3}$ (cubic, No. 206) and $Pnma$ (orthorhombic, No. 62), respectively. Ionic f^n terms^{16,17} and calculated values of $f''(\text{O}) = 0.06$ and $f''(\text{O}) = 0.04$ electrons per atom were used in the least-squares refinement of $f'(\text{Eu})$ and $f''(\text{Eu})$, the fractional coordinates and the usual profile parameters. (Anomalous scattering terms were calculated using the program FPRIME¹⁸.) A poor

fit to the Eu_3O_4 peaks was obtained (the residual for the integrated Eu_3O_4 diffraction peak intensities (defined in ref. 13), R_i , was 20.1%) when overall $f'(\text{Eu})$ and $f''(\text{Eu})$ values were used for all of the cation sites in Eu_2O_3 and Eu_3O_4 . To determine the valence distribution in Eu_3O_4 assuming no prior knowledge, I refined f' and f'' for one of the cation sites independent of the overall values for the other two, which were also varied. I did this for all three sites in turn and obtained R_i values of 18.5, 17.8 and 11.8% for the anomalous terms of sites Eu(1), Eu(2) and Eu(3) respectively. The significant improvement of the latter fit demonstrates that the Eu(3) site is electronically different from the other two. Table 1 shows the results of this refinement and the observed, calculated and difference profiles are given in Fig. 1.

To relate the refined f' and f'' values to the europium oxidation state, I used the reported results² for Gd^{3+} and Sm^{3+} to construct average curves (Fig. 2) for the variation of the anomalous scattering with the difference (ΔE) between the beam (E_λ) and absorption edge (E_{edge}) energies, $\Delta E = E_\lambda - E_{\text{edge}}$, where I have chosen the position of the minimum f' value for E_{edge} . These graphs should provide a good description of the f' and f'' variations for europium within the estimated standard deviations (e.s.d.s) in my refined values. Although ΔE is not a single-valued function of either f' or f'' , these two quantities change in very different ways around the edge position. Hence, the pair of anomalous scattering coefficients can be used to determine ΔE unambiguously and with reasonable precision, for small ΔE values, despite the large e.s.d.s in f' and f'' . I calculated energy shifts of −2.3(10) eV for the Eu(1) and (2) sites and 4.2(8) eV for Eu(3), as shown in Fig. 2. This gives the absolute positions

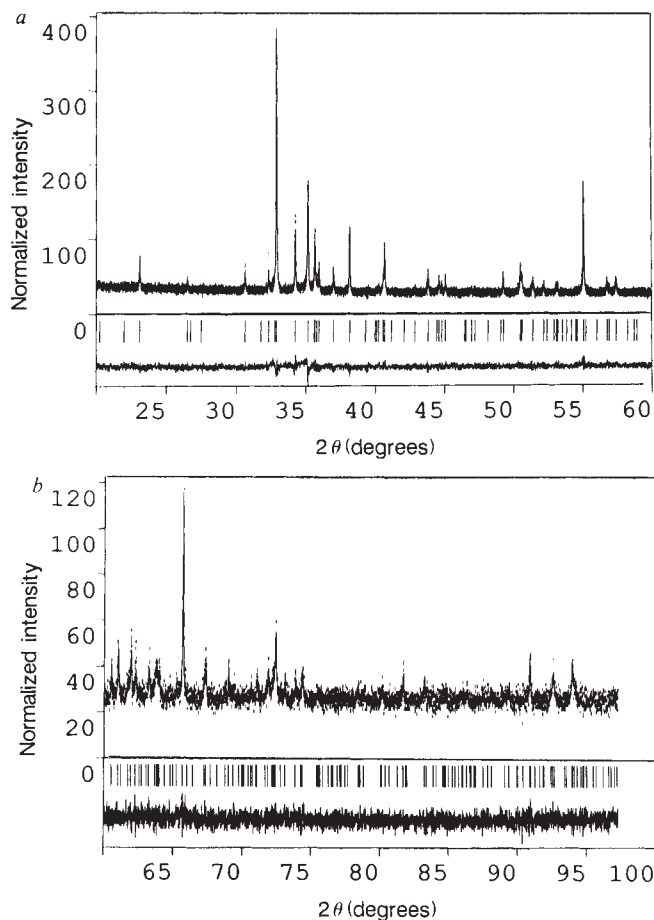


FIG. 1 Observed (ticks), calculated (full line), and difference synchrotron X-ray diffraction profiles for the $\text{Eu}_3\text{O}_4/\text{Eu}_2\text{O}_3$ sample, with reflection positions marked. Normalized intensities are in counts per s per 100-mA ring current.

of the L_{III} edges as 6,979.2(10) eV for Eu(1) and Eu(2) and 6,972(8) eV for Eu(3). The chemical shift of 6.5(13) eV between these values is in excellent agreement with the difference of 6.4(3) eV between the energies determined from the X-ray absorption spectra of Eu^{3+} in Eu_2O_3 at 6,976.1(2) eV and Eu^{2+} in EuO at 6,969.7(2) eV (ref. 19). The cause of the systematic error of 3 eV between the diffraction and spectroscopic results is unclear; it may arise from the different methods used to estimate the edge position.

The X-ray diffraction pattern of the europium samarium oxides sample was fitted by refining the structures of $\text{B-Sm}_2\text{O}_3$ (ref. 20) in $C2/m$ (monoclinic, No. 12) and the new Eu_3O_4 -type phase. The $\text{Eu}^{3+}/\text{Sm}^{3+}$ occupancies of the M(1) and M(2) sites and the $\text{Eu}^{2+}/\text{Sm}^{2+}$ distribution of the M(3) site in the latter compound were all refined independently using the previously determined anomalous scattering terms for Eu^{2+} and Eu^{3+} and calculated (using FPRIME¹⁸) values of $f'(\text{Sm}) = -10.32$ and $f''(\text{Sm}) = 9.80$ electrons per atom. The results in Table 1 clearly show that the former two sites are occupied only by Sm^{3+} whereas the M(3) site is occupied only by Eu^{2+} . The composition of this new phase is thus $\text{Eu}^{2+}(\text{Sm}^{3+})_2\text{O}_4$. This is consistent with the redox chemistries of these metals, as Eu^{3+} is more easily reduced to the +2 state than is Sm^{3+} .

To my knowledge, it has not previously been demonstrated that the X-ray absorption edge energies of crystallographically distinct sites can be determined with sufficient precision to distinguish directly between different valence states. This is not possible using normal X-ray scattering or neutron diffraction, except when applied to some magnetically ordered, mixed-valent transition metal compounds. Resonant X-ray diffraction experiments will be useful in studying ordered electronic states, especially when the coordination geometries are similar. Information concerning electronegativity²¹, spin state and phenomena such as charge density waves should also be accessible. The variation of the f' and f'' curves with these factors may be required to interpret the results of such experiments,

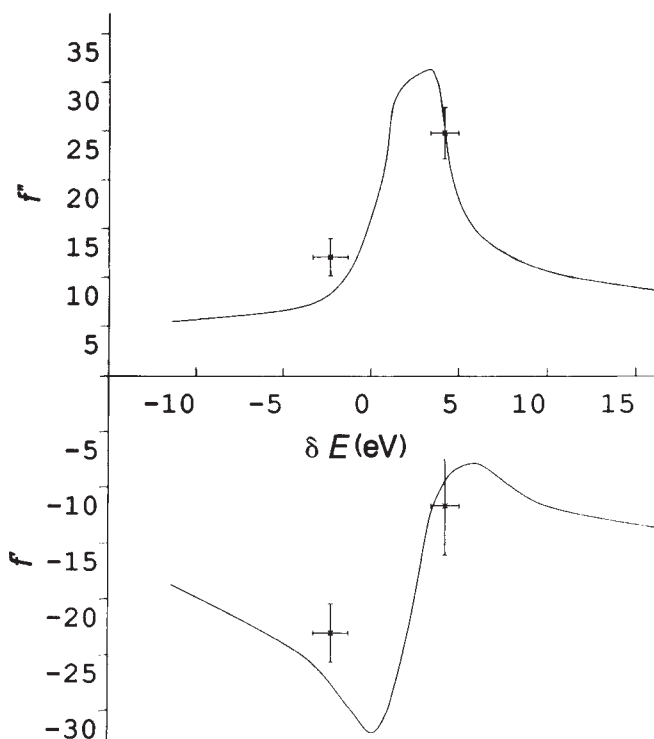


FIG. 2 Curves for the energy variation of f' and f'' (electrons per atom) around the $\text{Ln } L_{III}$ absorption edge ($\text{Ln}=\text{Gd-Sm}$). The points show the refined values for Eu^{2+} and Eu^{3+} in Eu_3O_4 .

TABLE 1 Results of the Rietveld refinements of Eu_3O_4 and EuSm_2O_4 in space group $Pnma$

R-factors (%)*					
	R_i	R_{wp}	R_p	R_{ex}	S_p^2 †
Eu_3O_4	11.8	9.3	7.4	8.6	1.17
EuSm_2O_4	13.5	10.2	8.1	9.0	1.27
Cell parameters (Å)					
	a		b	c	
Eu_3O_4	10.1050 (3)		3.5010 (1)	12.0675 (4)	
EuSm_2O_4	10.0955 (3)		3.5214 (1)	12.0986 (3)	
Refined atomic parameters (all atoms lie in the $y=3/4$ plane)					
Eu_3O_4					
Site	x	z	f' (electrons per atom)	f'' (electrons per atom)	
Eu(1)	0.4241 (17)	0.6153 (11)	-22.9 (26)	11.9 (20)	
Eu(2)	0.4101 (16)	0.1114 (11)	-22.9 (26)	11.9 (20)	
Eu(3)	0.2472 (11)	0.3549 (7)	-11.4 (45)	24.5 (27)	
O(1)	0.227 (8)	0.674 (6)			
O(2)	0.127 (7)	0.973 (6)			
O(3)	-0.001 (7)	0.219 (6)			
O(4)	0.451 (8)	0.911 (6)			
EuSm_2O_4					
Site	x	z	Site occupation factors (%)		
Eu(1)	0.4291 (14)	0.6167 (8)	$\text{Eu}^{3+} = -22(8)$	$\text{Sm}^{3+} = 122(8)$	
Eu(2)	0.4089 (13)	0.1104 (8)	$\text{Eu}^{3+} = -15(9)$	$\text{Sm}^{3+} = 115(8)$	
Eu(3)	0.2458 (12)	0.3539 (7)	$\text{Eu}^{2+} = 98(9)$	$\text{Sm}^{2+} = 2(9)$	
O(1)	0.213 (8)	0.680 (6)			
O(2)	0.130 (6)	0.992 (7)			
O(3)	-0.010 (7)	0.203 (5)			
O(4)	0.436(9)	0.919 (5)			

Fixed overall isotropic temperature factor $= 0.10 \text{ Å}^2$. The estimated standard deviations are given in parentheses.

^{*} The residuals are defined in ref. 13.

[†] The profile goodness-of-fit index is given by $S_p^2 = (R_{wp}/R_{ex})^2$.

and suitable curves can be derived in a straightforward manner from the X-ray absorption spectra of reference materials²².

The anomalous scattering terms also modify the scattering factor sufficiently to enable the resonant element to be distinguished even from its immediate neighbours in site-distribution refinements. This will provide a valuable alternative to using powder neutron diffraction data for such studies; neighbouring elements do not always have different neutron scattering lengths and it is difficult to collect an adequate diffraction pattern from a material containing nuclei with large incoherent scattering or absorption cross-sections, such as those of europium and samarium. It may also be possible to obtain very precise cation site distributions from a combined refinement with powder neutron and one or more resonant X-ray profiles (A. Williams *et al.*; and G. H. Kwei *et al.*, unpublished data).

Resonant X-ray diffraction experiments should be applicable to all elements that have an absorption edge at a convenient diffraction wavelength. The K edges of the elements from calcium to Zirconium and the L and M edges of elements heavier than silver all fall within a range of accessible, 'convenient' wavelengths of 0.7–3.1 Å. Although absorption by the sample is high in such experiments, which gives rise to a large background and lowers the precision in the refined atomic parameters, a flat-plate specimen makes absorption corrections unnecessary. Careful corrections would have to be made for other sample geometries. Continued advances in intense synchrotron X-ray sources and insertion devices should, in the future, enable good statistics to be obtained in reasonable counting times even for highly absorbing crystals or powders. □

Received 8 September; accepted 22 November 1989.

1. Attfield, J. P., Sleight, A. W. & Cheetham, A. K. *Nature* **322**, 620–622 (1986).
2. Templeton, L. K., Templeton, D. H., Phizackerley, R. P. & Hodgson, K. O. *Acta crystallogr.* **A38**, 74–78 (1982).

3. Karle, J. *Phys. Today* **42**(6), 22–29 (1989).
4. Flack, H. D. *Acta crystallogr.* **A39**, 876–881 (1983).
5. Will, G., Masciocchi, N., Hart, M. & Parrish, W. *Acta crystallogr.* **A43**, 677–683 (1987).
6. Moroney, L. M., Thompson, P. & Cox, D. E. *J. appl. Crystallogr.* **21**, 206–208 (1988).
7. Otsuka, K., Jinno, K. & Morikawa, A. *J. Catal.* **100**, 353–359 (1986).
8. Campbell, K. D., Zhang, H. & Lunsford, J. H. *J. phys. Chem.* **92**, 750–753 (1988).
9. Rau, R. C. *Acta crystallogr.* **20**, 716–723 (1966).
10. Rau, R. C. in *Rare Earth Research Vol. 2*, (ed. Vorres, K. S.) 117 (Gordon & Breach, New York, 1964).
11. *Nat. Bur. Stand. Monogr.* **25**, 20, 50 (1984).
12. Perakis, N. & Kern, F. C. *r. hebd. Séanc. Acad. Sci. Paris* **275B**, 677–680 (1972).
13. Rietveld, H. M. *J. appl. Crystallogr.* **2**, 65–71 (1969).
14. Wiles, D. B. & Young, R. A. *J. appl. Crystallogr.* **15**, 430–438 (1982).
15. Pauling, L. & Shappell, M. D. *Z. Kristallogr.* **75**, 128–142 (1930).
16. *International Tables for X-ray Crystallography Vol. 4*, 99–101 (Kynoch, Birmingham, 1974).
17. Tokonami, M. *Acta crystallogr.* **19**, 486 (1965).
18. Cromer, D. T. *J. appl. Crystallogr.* **16**, 437 (1983).
19. Finkel'shtein, I. D. et al. *Bull. Acad. Sci. USSR* **38**(3), 195–198 (1974).
20. Cromer, D. T. *J. phys. Chem.* **61**, 753–755 (1957).
21. Ghatikar, M. N. & Padalia, B. D. *J. Phys.* **C11**, 1941–1955 (1978).
22. Templeton, L. K. & Templeton, D. H. *J. appl. Crystallogr.* **21**, 558–561 (1988).

ACKNOWLEDGEMENTS. I thank A. Fitch and K. P. J. O'Reilly for assistance during the data collection, the SRS Powder Diffraction Service for the provision of beam time, D. J. Watkin for comments on the manuscript, and Christ Church, Oxford for a Junior Research Fellowship.

Sensitivity of the Earth's radiation budget to changes in low clouds

A. Slingo

Climate and Global Dynamics Division, National Center for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307-3000, USA

VARIOUS mechanisms have been suggested whereby clouds might take part in or initiate climate change, including changes in cloud amounts, liquid-water paths and droplet sizes^{1–11}. Previous studies of the sensitivity of the Earth's radiation budget to cloud liquid-water path and droplet size were made with one-dimensional or even simpler models^{1–6,12,13}, which cannot represent the real cloud distribution. Here I present the results of a study that uses a three-dimensional general circulation model, which should give more reliable estimates. The top-of-atmosphere radiative forcing by doubled carbon dioxide concentrations can be balanced by modest relative increases of ~15–20% in the amount of low clouds and 20–35% in liquid-water path, and by decreases of 15–20% in mean drop radius (depending on the version of the model). This indicates that a minimum relative accuracy of ~5% is needed, both to simulate these quantities in climate models and to estimate climate response by monitoring them over extended periods from satellite platforms.

Results from the Earth Radiation Budget Experiment¹⁴ for April 1985 show that the net effect of clouds at the top of the atmosphere (TOA) is a cooling of 13 W m⁻². Because the change in the TOA net radiation that is due to an instantaneous doubling of the carbon dioxide concentration is only ~2.5 W m⁻² (see later), changes in cloudiness could contribute significantly to climate change.

I focus on the effect of low clouds (within 2 km of the surface) because they strongly influence the planetary albedo^{14,15}, have been invoked in a discussion of the climate change brought about by variations in particle size¹ and may be sensitive to anthropogenic influences¹⁶.

I consider this problem in the framework of a global climate feedback analysis^{17,18}. In equilibrium, the globally and annually averaged net radiation at the top of the atmosphere N (the difference between the absorbed shortwave radiation Q and outgoing longwave radiation F) should be zero, $N = Q - F = 0$. Climate change occurs when the system is driven out of radiative equilibrium by a perturbation in either (or both) Q and F , $\Delta N = \Delta Q - \Delta F$. The response of the surface temperature T_s may be expressed in terms of a climate sensitivity parameter λ , $\Delta T_s = \lambda \Delta N$. The parameter λ can be expressed in terms of the

TABLE 1 Brief specifications of the versions of the model

Parameter	Version of CCM		
	1	2	3
Spectral truncation*	R15	R15	T42
Cloud prediction scheme†	Standard	ECMWF	ECMWF
Liquid water path profile (Fig. 1)	A	B	B

* The approximate horizontal latitude $\times$ longitude resolution in degrees is 4.5×7.5 for R15 and 2.8×2.8 for T42.

† Described in refs 21 and 23 (Standard) and ref. 24 (ECMWF).

'gain' G_0 of the climate system in the absence of feedbacks, plus a feedback factor f : $\lambda = G_0/(1-f)$. Equilibrium is restored because $F = \varepsilon_a \sigma T_s^4$ (where ε_a is the mean atmospheric emissivity and σ is the Stefan-Boltzmann constant), which may be differentiated to give $G_0 = T_s/4F = 0.3 \text{ K m}^2 \text{ W}^{-1}$. For $f = 0$, $\lambda = G_0$ and there is no feedback. For $f > 0$ there is amplification of G_0 and hence positive feedback, whereas for $f < 0$ there is damping of G_0 and negative feedback.

Because many feedbacks may be operating simultaneously,

$$f = G_0 \sum_j \frac{\partial N}{\partial I_j} \frac{\partial I_j}{\partial T_s} \quad (1)$$

This conveniently breaks down the calculation of f into three more manageable components: the gain G_0 in the absence of feedbacks, how the net radiation varies with internal variable I_j and how internal variable I_j varies with T_s . Note that the sign of the feedback depends on the sign of both $\partial N/\partial I_j$ and $\partial I_j/\partial T_s$. The latter term is the most difficult to determine and the most uncertain component of the feedback. If I_j represents clouds then ∂N is the change in the cloud radiative forcing^{14,19} (the effect of clouds on the present distribution of radiative heating), which is thus a component of the cloud radiative feedback²⁰.

I concentrate here on estimates of $\partial N/\partial I_j$ for low clouds, to determine the sensitivity of the Earth's radiation budget (ERB) to cloud amount, liquid-water path and droplet size. I used the 12-layer NCAR Community Climate Model (CCM1)²¹, modified to include a new parameterization²² for the shortwave radiative properties of water clouds. In this parameterization I have separated the effects of the cloud liquid-water path (LWP) and the droplet equivalent radius, r_e . Here LWP and r_e are prescribed (see below), but they will be variables in future work. These parameters are important because their ratio determines the cloud albedo through the optical depth τ :

$$\tau \approx \frac{3}{2} \frac{\text{LWP}}{r_e} \quad (2)$$

where LWP is in g m⁻² and r_e is in μm . For consistency, the liquid-water path was also used to calculate the cloud longwave emissivity ε , $\varepsilon = 1 - \exp(-0.1 \text{ LWP})$.

I have used three versions of the model (Table 1). Version 1 uses the standard R15 resolution and cloud prediction

TABLE 2 Changes in the Earth's radiation budget at the top of the atmosphere due to an instantaneous CO₂ doubling

Time of year	Radiation budget quantity (W m ⁻²)	Version of CCM		
		1	2	3
January	Absorbed shortwave	0.15	0.15	0.15
	Outgoing longwave	-2.20	-2.27	-2.34
	Net radiation	2.35	2.41	2.49
July	Absorbed shortwave	0.15	0.15	0.15
	Outgoing longwave	-2.41	-2.47	-2.54
	Net radiation	2.56	2.62	2.69

scheme^{21,23}. Version 2 is the same except that the cloud scheme was replaced by that used in the ECMWF medium-range forecast model²⁴. Version 3 is the same as 2, with the resolution increased to T42. For each version, integrations were made using prescribed sea surface temperatures for perpetual January and July conditions and the results averaged over 500 days of simulation. I adjusted the parameters of the ECMWF cloud scheme used in versions 2 and 3 so that the globally averaged total cloud amount was close to 50%. The equivalent radius was assumed to be 10 μm for all clouds and the LWP profiles were also prescribed (Fig. 1). To obtain the profiles I used simple interpolation between values that give reasonable radiative properties for both low and high clouds. The shortwave optical depth of ~ 10 for lower tropospheric clouds (Fig. 1) is in agreement with the International Satellite Cloud Climatology Project (ISCCP) data²⁵. For version 1, the simulations of the cloud and ERB distributions are similar to those from the standard model²³. For versions 2 and 3, the simulations are much improved by the new cloud scheme. In particular, the deficiencies in the standard model at T42 resolution²⁶ are removed in version 3, which is regarded as providing the most reliable sensitivities. In addition, the globally averaged longwave, shortwave and net cloud radiative forcings are also realistic; the largest differences compared with ERBE¹⁴ are 7, 5 and 3 W m^{-2} for the three versions, respectively. Details of the modifications to the model and of the effect on model simulations will be presented elsewhere (A. Slingo and J. M. Slingo, manuscript in preparation).

I obtained the estimates of $\partial N/\partial I_j$ as follows. A one-time-step integration was performed and the globally averaged radiation budget obtained. This integration regenerated exactly the control calculation at a given time step. I could then determine the effect of changing various parameters one by one and calculate the partial derivatives from the differences. I used about twenty samples through each of the six control integrations (three versions of the model for both January and July) to ensure statistically reliable results. The parameters chosen were carbon dioxide concentration and low-cloud amount, LWP and r_e . The cloud parameters were multiplied by various factors to bracket the ΔN produced by doubling the amount of carbon dioxide.

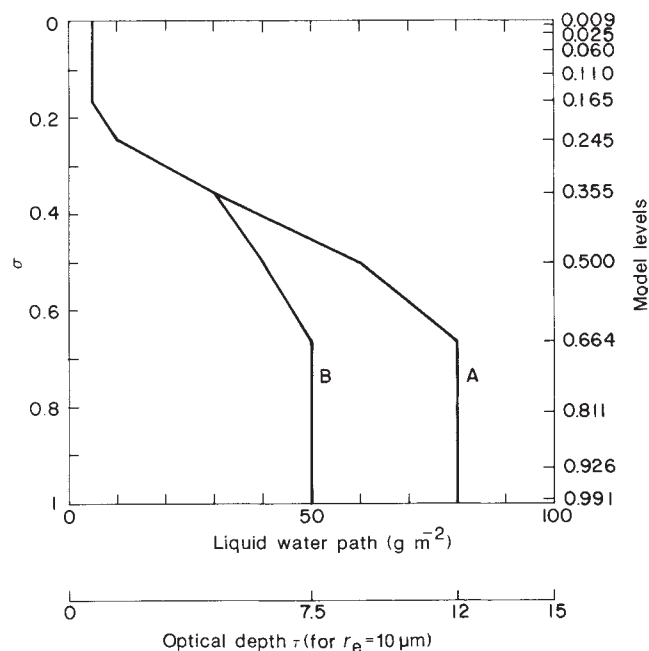


FIG. 1 Vertical profiles of cloud liquid water path used in the CCM integrations. The vertical coordinate is σ (pressure divided by its surface value) and the 12 model levels are shown. The approximate scale for cloud optical depth derived from equation (2) is shown at the bottom.

TABLE 3 Factors to counteract net radiation changes

Time of year	Quantity	Version of CCM		
		1	2	3
January	Amount	1.19	1.14	1.14
	LWP	1.31	1.20	1.21
	r_e/n	0.81/1.90	0.86/1.60	0.85/1.65
	$r_e^\dagger/n^\dagger$	0.77/2.16	0.82/1.81	0.81/1.89
July	Amount	1.20	1.15	1.15
	LWP	1.35	1.21	1.23
	r_e/n	0.79/2.02	0.85/1.65	0.83/1.72
	$r_e^\dagger/n^\dagger$	0.71/2.85	0.78/2.10	0.75/2.35

Factors by which the global low-cloud amount, liquid water path, droplet equivalent radius and number concentration* need to be multiplied to counteract the changes in net radiation shown in Table 2.

* Derived from the r_e factor (before rounding), as $n \propto r_e^{-3}$ at constant LWP.

† For changes in r_e imposed only over the open (ice-free) oceans.

Because the cloud amount in many areas was already close to 100%, I reduced the amounts by various factors and assumed that the sensitivity was the same as if the amounts had been increased.

Table 2 shows the change in the ERB when the amount of carbon dioxide is doubled. There is only a weak dependence on the version of the model used. Although $\partial N/\partial I_j$ can be found from the information given here, Table 3 shows the results more simply as the factor by which the low-cloud quantity must be multiplied to produce a change in the net radiation equal and opposite to that obtained by doubling the carbon dioxide (Table 2). For example, the low-cloud amount in January for version 3 must be multiplied by 1.14, a relative increase of 14%.

There is a significant dependence of the results in Table 3 on the version of the model, mainly because of the different cloud schemes. Version 1 is the least sensitive to low clouds, requiring larger changes to balance the carbon dioxide forcing. Considering first the low-cloud amount, the carbon dioxide forcing can be balanced by relative increases of between 14% and 20%. These correspond to absolute increases of between 3.5% and 5%, as the model's globally averaged low-cloud cover is $\sim 25\%$. Previous estimates range from a 10% relative increase¹³ to a 4% absolute increase in the area of the globe covered by low-level marine stratus¹⁵ (corresponding to a relative increase of $\sim 17\%$, assuming that 23.7% of the Earth's surface is covered by such clouds¹). For liquid-water path, the variation is between 20% and 35%. Charlson's results⁴ for the sensitivity of the absorbed shortwave to the liquid-water path for an initial value of 50 g m^{-2} (his Tables 1 and 2) imply a figure of $\sim 40\%$. Finally, the assumed drop radius of 10 μm must be reduced to between 7.9 and 8.6 μm to balance the carbon dioxide forcing. Assuming that the liquid-water path does not change, these values correspond to increases in drop number concentrations of between 60% and 102% (Table 3). This may be compared with results⁵ that indicate that the carbon dioxide forcing may be offset by a 41% increase in the number of drops in clouds at all levels (not just low). If only low clouds over the open oceans are considered, the present figures increase to 81–185% (Table 3), which spans the figure of $\sim 100\%$ found by Charlson *et al.*¹

These results are thus in broad agreement with those from simpler models, although they underline the sensitivity of climate-change projections to model details, particularly in the cloud prediction scheme. By normalizing to the carbon dioxide forcing, I conclude that a minimum relative accuracy of $\sim 5\%$ is needed, both to model these cloud quantities and monitor them from satellite measurements, if their contribution to climate change is to be determined. For low-cloud amount, this corresponds to an accuracy requirement of $\sim 1\%$ in the absolute amount. This is a stringent requirement, and emphasizes the advances needed both to model and monitor the effects of clouds on climate. □

Received 25 September; accepted 16 November 1989.

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
2. Schneider, S. H. *J. Atmos. Sci.* **29**, 1413–1422 (1972).
3. Paltridge, G. W. *Q. J. R. Met. Soc.* **106**, 895–899 (1980).
4. Charlson, T. P. *Tellus* **34**, 245–254 (1982).
5. Twomey, S. A., Piepgrass, M. & Wolfe, T. L. *Tellus* **36B**, 356–366 (1984).
6. Somerville, R. C. J. & Remer, L. A. *J. Geophys. Res.* **89**, 9668–9672 (1984).
7. Schlesinger, M. E. & Mitchell, J. F. B. *Rev. Geophys.* **25**, 760–798 (1987).
8. Wetherald, R. T. & Manabe, S. *J. Atmos. Sci.* **45**, 1397–1415 (1988).
9. Schwartz, S. E. *Nature* **336**, 441–445 (1988).
10. Wigley, T. M. L. *Nature* **339**, 365–367 (1989).
11. Mitchell, J. F. B., Senior, C. A. & Ingram, W. J. *Nature* **341**, 132–134 (1989).
12. Stephens, G. L. & Webster, P. J. *J. Atmos. Sci.* **38**, 235–247 (1981).
13. Webster, P. J. & Stephens, G. L. in *The Global Climate* (ed. Houghton, J. T.) 63–78 (Cambridge University Press, 1984).
14. Ramanathan, V. *et al. Science* **243**, 57–63 (1989).
15. Randall, D. A. *et al. Bull. Am. Met. Soc.* **65**, 1290–1301 (1984).
16. Coakley, J. A., Bernstein, R. L. & Durkee, P. A. *Science* **237**, 1020–1022 (1987).
17. Schlesinger, M. E. *Nature* **336**, 315–316 (1988).
18. Schlesinger, M. E. in *Physically-Based Modelling and Simulation of Climate and Climatic Change Part 2* (ed. Schlesinger, M. E.) 653–735 (Kluwer, Dordrecht, 1988).
19. Slingo, A. & Slingo, J. M. *Q. J. R. Met. Soc.* **114**, 1027–1062 (1988).
20. Cess, R. D. *et al. Science* **245**, 513–516 (1989).
21. Williamson, D. L. *et al. Description of NCAR Community Climate Model (CCM1) NCAR Technical Note NCAR/TN-285 + STR* (Nat. Cent. Atmos. Res. Boulder, 1987).
22. Slingo, A. *J. Atmos. Sci.* **46**, 1419–1427 (1989).
23. Ramanathan, V. *et al. J. Atmos. Sci.* **40**, 605–630 (1983).
24. Slingo, J. M. *Q. J. R. Met. Soc.* **113**, 899–927 (1987).
25. *Results from the International Satellite Cloud Climatology Project* (ISCCP)* (World Climate Research Programme Radiation Projects Office, Geneva, 1989).
26. Kiehl, J. T. & Ramanathan, V. *J. Geophys. Res.* (in the press).

ACKNOWLEDGEMENTS. The National Center for Atmospheric Research is sponsored by the NSF.

Climate-induced changes in forest disturbance and vegetation

Jonathan T. Overpeck*, David Rind†
& Richard Goldberg‡

* Lamont–Doherty Geological Observatory of Columbia University,
Palisades, New York 10964, USA

† NASA Goddard Institute for Space Studies, 2880 Broadway,
New York, New York 10025, USA

‡ Institute for Global Habitability, Columbia University,
New York 10027, USA

RECENT concern over the ecological effects of future trace-gas-induced climate change has accelerated efforts to understand and quantify climate-induced vegetation change^{1–9}. Here we discuss new and published climate-model results indicating that global warming favours increased rates of forest disturbance, as a result of weather more likely to cause forest fires (drought, wind and natural ignition sources), convective wind storms, coastal flooding and hurricanes. New sensitivity tests carried out with a vegetation model indicate that climate-induced increases in disturbance could, in turn, significantly alter the total biomass and compositional response of forests to future warming. An increase in disturbance frequency is also likely to increase the rate at which natural vegetation responds to future climate change. Our results reinforce the hypothesis⁶ that forests could be significantly altered by the first part of the next century. Our modelling also confirms the potential utility of selected time series of fossil pollen data for investigating the poorly understood natural patterns of century-scale climate variability.

We wish to highlight the need to incorporate realistic disturbance regimes in model assessments of future vegetation change. Higher rates of disturbance can increase the rate at which forests are opened up for new trees that are better adapted to modified climate conditions. Disturbance thus acts to short-circuit the slower process of forest-gap initiation and overturn through old-age mortality^{10,11}. Increased rates of forest disturbance also increase the relative proportion of early successional (rapid-growing, shade-intolerant) and disturbance-adapted trees on the landscape. The form and rate of climate-induced change in vegetation can thus depend on the frequency of disturbance^{10,11}.

In northern mid-latitude forests, the chief forms of non-anthropogenic catastrophic disturbance are fire and catastrophic windthrow^{12,13}. The incidence of these events is proportional to the incidence of 'disturbance weather', primarily summer/autumn drought and thunderstorms (lightning and wind) for fire¹⁴ and large-scale thunderstorms for windthrow¹². To a lesser degree, disturbance weather also includes tornadoes, hurricanes and floods. Climate-model results indicate that warmer climates tend to be characterized by more frequent disturbance weather for several reasons. Latitudinal temperature contrasts and atmospheric eddy energy tend to be weakened as climate warms, thus decreasing temperature variability on all timescales¹⁵. This may lead to longer runs of consecutive hot days without cool breaks during future growing seasons. In addition, the increased moisture-holding capacity of warmer atmospheres should be associated with greater precipitation variability¹⁵. As climates become warmer, temperature-driven increases in potential evapotranspiration will probably exceed increases in precipitation at the warmer low and mid-latitudes^{16,17}. All of these factors combine to make the probability of summer/autumn drought, and thus fire, more likely during warmer periods^{15–18}.

Model results also indicate that higher temperatures and greater atmosphere moisture loading produce significant increases in penetrating convection (Fig. 1). This increase implies a significant change in the intensity and frequency of large thunderstorms. Several studies using general circulation model (GCM) results also indicate that tropical storms and hurricanes are likely to be strengthened under warm climate conditions^{19,20}. Climate theory and model results therefore support recent analyses of palaeoecological fire-frequency records^{21–23}, which offer a clear indication that the magnitude and frequency of disturbance weather could be significantly higher during periods of hemispheric warmth than during cooler times.

We used the mixed-species (72 total), mixed-age stochastic stand-simulation model FORENA to simulate the effects of climate change on four forest types in eastern North America (Table 1). FORENA is a widely used and tested model that simulates the processes controlling the establishment, growth and death of trees on a 1/12-ha plot^{5,24}. We modified the program so that we could specify the probability of catastrophic disturbance or death of all trees on a given plot. We also specified that

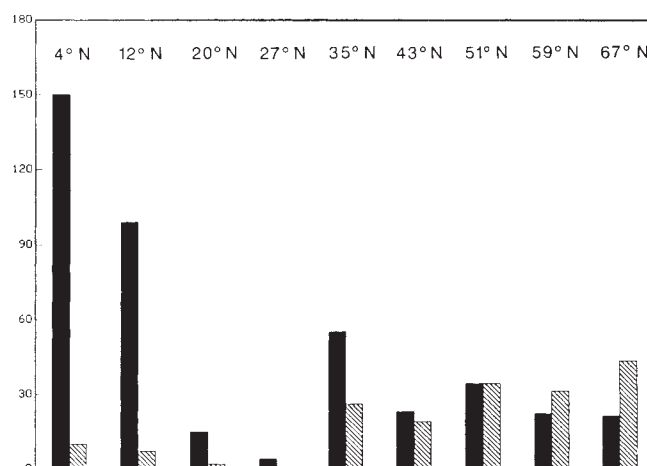


FIG. 1 Increase in annual average penetrative mass flux that is due to moist convection through the 300-mbar level as a function of latitude in an equilibrium $2 \times \text{CO}_2$ climate-change simulation²⁵. Values represent mass flux × height (solid bars in $10^{12} \text{ kg m s}^{-1}$), and percentage increase in mass flux (striped bars). An increase in mass flux to higher levels in the atmosphere is likely in a world warmed by trace gases. This increase will probably translate into more intense thunderstorms, wind, lightning activity (C. Price and D. Rind, manuscript in preparation) and forest disturbance at mid- to high latitudes.

TABLE 1 Sites at which forest growth was simulated

State or province	Latitude	Longitude	Dominant modern-day forest constituents
N Wisconsin	45.0°N	90.0°W	<i>Pinus</i> , <i>Acer</i> , <i>Quercus</i>
S Quebec	47.9°N	75.0°W	<i>Abies</i> , <i>Picea</i> , <i>Betula</i>
NE Michigan	44.8°N	83.6°W	<i>Pinus</i> , <i>Quercus</i> , <i>Picea</i>
S Illinois	39.0°N	89.0°W	<i>Quercus</i> , <i>Carya</i>

a plot had to remain undisturbed for a period of 20 regeneration years before another disturbance could occur. We carried out two types of sensitivity experiments. In the first set (the 'step-function' experiments), we compared pairs of perturbation (with a disturbance increase) and nonperturbation (without a change in disturbance) runs. In both of these runs, simulated forest was grown from bare ground under present-day climate⁴ for 800 years to characterize the natural variability of the simulated forest. At year 800, a single climate variable was then changed in a single step to a new mean climate and the simulation run for a further 400 years. In each perturbation experiment, we changed the probability of catastrophic disturbance from 0.00 to 0.01 (equivalent to a realistic frequency of about 1 plot-destroying fire every 115 years, when the 20-year regeneration period is factored in) at year 800; a 0.00 probability was maintained throughout each nonperturbation run.

We examined three types of climate change separately in our step-function experiments: (1) a 1°C temperature increase; (2) a 2°C increase; (3) a 15% decrease in precipitation. Changes of these magnitudes could easily have occurred within the late Quaternary and could occur in the future. In our second set of experiments (the 'transient' experiments), we also grew forest from bare ground under observed climate for 800 years. From year 800 to year 900, we linearly changed the mean monthly climate (temperature and precipitation) to simulated equilibrium $2 \times \text{CO}_2$ (twofold increase in CO_2) levels^{4,25} and then held the climate constant until year 1600. Our rate of change was slightly less than the scenario 'A' transient response modelled by Hansen *et al.*²⁵. As before, the nonperturbation runs contained no change in disturbance probability, whereas the perturbation experiments contained step-function increases (from 0.00 to 0.01) in the probability of disturbance at year 800. We used the same relatively drought-insensitive soil in all our runs⁵, and averaged the results from 40 random plots into a single time series for each model run.

Selected summary results from sites in the mixed conifer-hardwood forest of Wisconsin and the southern boreal forest of Quebec (Fig. 2) illustrate our results, which are supported by our other experiments, not shown here. An increase in forest disturbance will probably produce climate-induced vegetation change that is greater than, or equal to, the same climate-induced change in the absence of an altered disturbance regime. In many cases, this enhanced change is due to the rapid increases in the

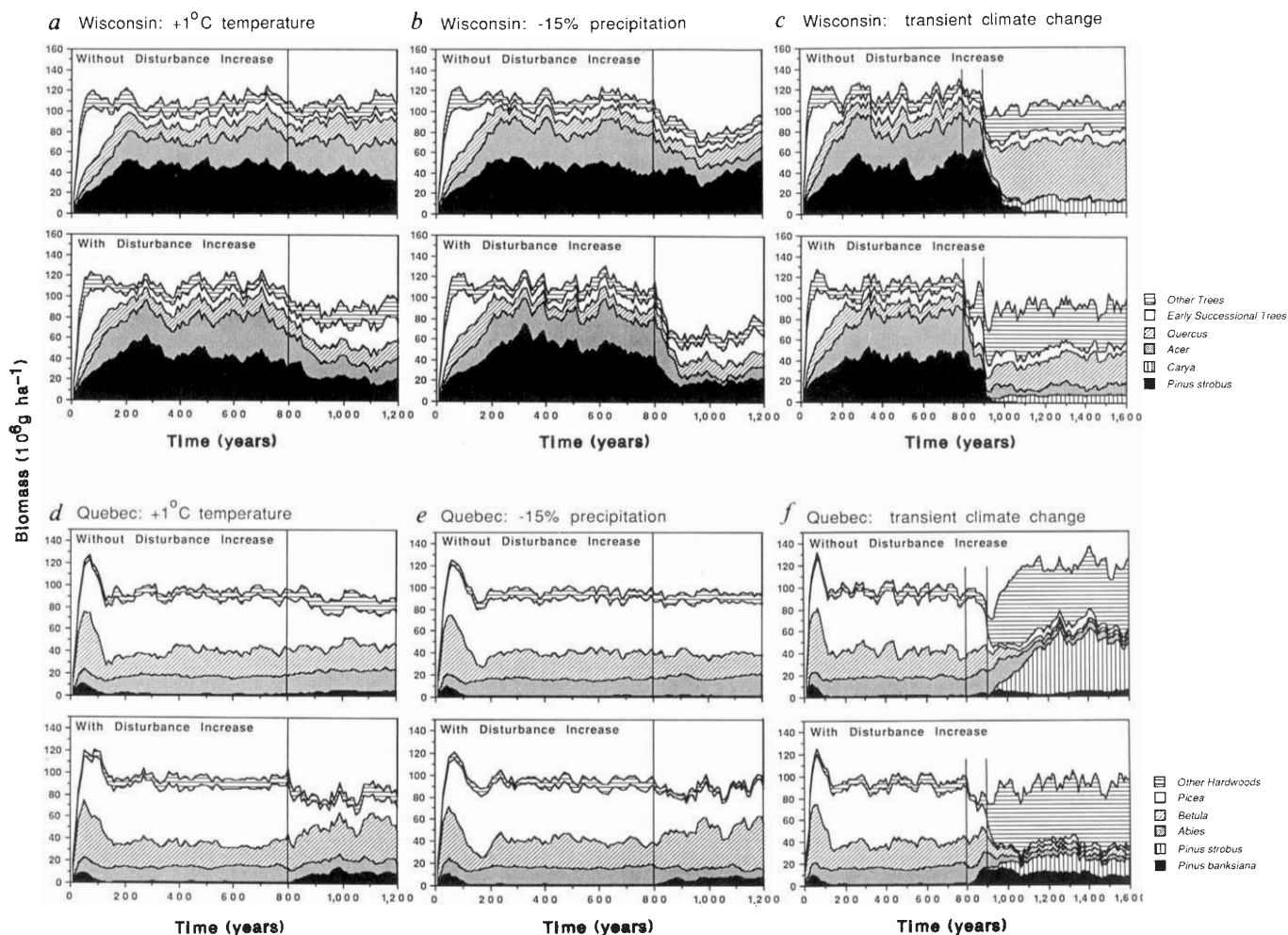


FIG. 2 Simulated changes in species composition in forests at two of the sites we investigated in eastern North America: a–c in Wisconsin, d–f in southern Quebec (see Table 1). Each plot consists of a nonperturbation

simulation (without a change in disturbance) and a perturbation (with a change in disturbance) simulation.

abundances of early successional species that occur in response to increased disturbance frequency. In some cases (such as Figs 2a, d and e), a step-function change in climate alone did not cause a significant change in simulated biomass, whereas the same climate change coupled with an increase in forest disturbance resulted in a significant effect on the forest.

Not only did altered disturbance regimes tend to cause significant changes in simulated forest composition, but they also tended to accelerate the rate of forest response to climate change. The increase in forest disturbance never led to a slower response. In the case of the transient experiments, vegetation change generally lagged behind the initiation of climate change by 50–100 years, in the absence of an increased frequency of disturbance. In these same transient nonperturbation runs, the simulated vegetation took at least 200–250 years to achieve equilibrium conditions under the $2 \times \text{CO}_2$ climate. By contrast, the presence of an increased rate of forest disturbance resulted in vegetation change that more closely tracked the initiation and timing of climate change. In all of the transient-perturbation experiments, the simulated vegetation took <180 years to regain equilibrium after the initiation of climate change at year 800.

Our results should not be construed as predictions for the future trace-gas-warmed world, but instead highlight the need for improved vegetation models (both on a regional and a global scale) that include realistic disturbance routines that are interactively linked to climate variability. Other factors (such as soil dynamics, seed sources, direct trace-gas effects, pollution, pests, pathogens, forest management and the accumulation of fire fuels) must be considered^{8,23}. Rapid climate change could speed tree mortality and forest dieback⁵, thus increasing the rate of fuel buildup and the probability of disturbance by fire.

Our results have immediate applicability for the future, but they also indicate that vegetation change may have closely tracked past climate change in the presence of disturbance regimes that were suitably rapid and sensitive to climate variability. This supports the careful use of fossil-pollen data in the examination of past century to millenium-scale climate change^{26–28}. Such studies should be important for assessing the rapid climate and vegetation changes that may occur in the future. Our results thus contrast somewhat with earlier forest-modelling results¹⁰, which indicated that pollen records may not be capable of resolving climate change within 100–150 years. This assertion is undoubtedly true for many sites with insensitive forest-disturbance regimes, but should not hinder the continued use of fossil pollen data in the study of the patterns and causes of century-scale climate and vegetation variability. □

Received 19 July; accepted 3 November 1989.

1. Bartlein, P. J., Prentice, I. C. & Webb, T. III *J. Biogeog.* **13**, 35–57 (1986).
2. Webb, T. III, Bartlein, P. J. & Kutzbach, J. E. in *The Geology of North America* Vol. K-3 (eds Ruddiman, W. F. & Wright, H. E. Jr) 447–462 (Geol. Soc. Am., Boulder, 1987).
3. COHMAP Members *Science* **241**, 1043–1052 (1988).
4. Overpeck, J. T. & Bartlein, P. J. in *U.S. Environmental Protection Agency Rep. EPA-230-05-89-54* Ch. 1, 1–32 (US EPA, Washington DC, 1989).
5. Solomon, A. M. *Oecologia* **68**, 567–579 (1986).
6. Botkin, D. B., Nisbet, R. A. & Reynolds, T. E. in *U.S. Environmental Protection Agency Rep. EPA-230-05-89-54* Ch. 2, 1–31 (US EPA, Washington DC, 1989).
7. Pastor, J. & Post, W. M. *Nature* **334**, 55–58 (1988).
8. Solomon, A. M. & West, D. C. in *U.S. Department of Energy, Rep. DOE/ER-0236*, 145–169 (US DOE, Washington DC, 1985).
9. Neilsen, R. P. *Science* **232**, 27–34 (1986).
10. Davis, M. B. & Botkin, D. B. *Quat. Res.* **23**, 327–340 (1985).
11. Green, D. G. *J. Biogeog.* **9**, 29–40 (1982).
12. Canham, C. D. & Loucks, O. L. *Ecology* **65**, 803–809 (1984).
13. Pickett, S. T. A. & White, P. S. *The Ecology of Natural Disturbance and Patch Dynamics* (Academic, Orlando, 1985).
14. Schroeder, M. J. & Buck, C. C. *U.S. Dept. Agric. Handbk* 360 (US Government Printing Office, Washington DC, 1970).
15. Rind, D., Goldberg, R. & Ruedy, R. *Clim. Change* **14**, 5–37 (1989).
16. Hansen, J. et al. in *Preparing for Climate Change, Proc. 2nd North American Conf. on Preparing for Climate Change* (Climate Institute, Washington DC, in the press).
17. Rind, D., Goldberg, R., Hansen, J., Rosenzweig, C. & Ruedy, R. *J. geophys. Res.* (submitted).
18. Manabe, S. & Wetherald, R. T. *J. Atmos. Sci.* **44**, 1211–1236 (1987).
19. Emanuel, K. A. *Nature* **326**, 483–485 (1987).
20. Hobgood, J. S. & Cerveny, R. S. *Nature* **333**, 243–245 (1988).
21. Winkler, M. G., Swain, A. M., & Kutzbach, J. E. *Quat. Res.* **25**, 235–250 (1986).
22. Gajewski, K. *Vegetatio* **68**, 179–190 (1987).
23. Clark, J. S. *Nature* **334**, 233–235 (1988).

24. Shugart, H. H. *A Theory of Forest Dynamics*. (Springer, New York, 1984).
25. Hansen, J. et al. in *Climate Processes and Climate Sensitivity* (eds Hansen, J. & Takahashi, T.) 130–163 (Am. geophys. Un., Washington, DC, 1984).
26. Bernabo, J. C. *Quat. Res.* **15**, 143–159 (1981).
27. Overpeck, J. T. in *Abrupt Climatic Change—Evidence and Implications* (eds Berger, W. H. & Labeyrie, L. D.) 137–143 (Reidel, Dordrecht, 1987).
28. Gajewski, K. *Quat. Res.* **29**, 255–262 (1988).

ACKNOWLEDGEMENTS. We thank A. Solomon and L. Tharp for providing the stand-simulation model (FORENA) and insights about its design and use. We also thank E. Cook, J. Cole, G. Jacoby, S. Miller, C. Price, and T. Webb III for discussions and comments on the manuscript. Funding was provided by the US Environmental Protection Agency and US National Science Foundation Division of Atmospheric Sciences (Climate Dynamics).

Evaluation of an acidification model with data from manipulated catchments in Norway

R. F. Wright*, B. J. Cosby†, M. B. Flaten* & J. O. Reuss‡

* Norwegian Institute for Water Research, Box 69 Korsvoll, 0808 Oslo, Norway

† School of Forestry and Environmental Sciences, Duke University, Durham, North Carolina 27706, USA

‡ Department of Agronomy, Colorado State University, Fort Collins, Colorado 80523, USA

INTERNATIONAL efforts to control the emissions of the acidifying compounds SO_2 and NO_x to the atmosphere, and to establish critical loads, are in part aimed at predicting the long-term response of soils and waters to changes in acid deposition. Estimation of future long-term trends in acidification requires the use of models, the strict verification of which requires years or decades to determine whether predictions match observed responses. Experimental manipulations with whole ecosystems provide data with which predictive models can be evaluated. The MAGIC model^{1,2} is a widely used, intermediate-complexity, process-oriented model for soil and water acidification in catchments. Here we follow up a preliminary application³ of the model to the first two years of data from the manipulated catchments of the RAIN (Reversible Acidification In Norway) project⁴. We use a more extensive four-year data set as inputs to a new version of the model. The four-year record provides a much more rigorous test of the predictive ability of the model because these data show major changes in runoff chemistry due to the experimental change in deposition. Sensitivity analysis shows that MAGIC is successful at predicting future acidification.

The RAIN project comprises two parallel large-scale manipulations in which acid deposition is experimentally changed at whole catchments⁴. At Sogndal in western Norway, a 7,200-m² pristine headwater catchment (SOG2) is artificially acidified by the addition of 70–100 meq m⁻² yr⁻¹ H_2SO_4 . At Risdalsheia in southern Norway, ambient acid precipitation is excluded by means of a 1,200-m² roof, and clean precipitation is applied beneath the roof. Two untreated catchments at each site provide reference data.

The processes incorporated in the MAGIC model include atmospheric deposition, sulphate adsorption, cation exchange, CO_2 dissolution, precipitation and dissolution of aluminium, chemical weathering, uptake and release of cations by vegetation and export in runoff. The model produces long-term reconstructions and predictions of soil and streamwater chemistry in response to given acid deposition scenarios^{1,2}. It uses an aggregated approach in two ways: (1) a myriad of chemical and biological processes active in catchments are modelled by a few readily described processes; (2) an averaged set of parameters is used to describe the soil properties within the catchment.

Whereas standard precipitation gauges provide adequate estimates of integrated inputs to catchments and the outputs in runoff are integrated at the weir, estimates of the soil parameters

characteristic of an entire catchment are more difficult to obtain. Key soil parameters used in the model include depth, bulk density, porosity, cation exchange capacity (measured at soil pH) and the fraction of exchange sites occupied by calcium, magnesium, sodium and potassium. We estimate these from soil data collected annually^{5,6}. Values are averaged both spatially and with depth at each catchment to obtain single values for each parameter. We use the mean value of data expressed as logarithms for square roots to compensate for skew (Table 1).

Data for precipitation and runoff chemistry at each site are given in ref. 4. Our calibration procedure involves reconstruction of historical soil and runoff chemistry from pre-acid deposition conditions (assumed to be about 140 years ago) to 1983, the year before treatment at the RAIN catchments began. Historical acidification is driven by acid deposition which we assume has been proportional to emissions of SO_2 in Europe⁷.

We use an optimization routine to calibrate the model for each catchment. This routine chooses rates for chemical weather-

ing and pre-acidification base saturation in the soils such that after 140 years of acid deposition the model gives the measured chemical composition of runoff and measured soil chemistry in 1983–1984, the year before treatment began.

The calibrated model predicts changes in runoff and soil chemistry as a result of the experimental manipulation of acid deposition at Sogndal and Risdalsheia. To better account for yearly and monthly variations in deposition volume and chemical composition, we use here a modified version of the MAGIC model in which the simulation, calibrated to 1983, is run using the actual measured monthly deposition inputs (including acid additions) for the four-year period 1984–1987. This monthly version also takes into account the buildup of snow during the winter and the spring melt. Simulated output is then compared to the chemical composition of the actual monthly runoff over the four years of treatment (Sogndal, November 1983–November 1987; Risdalsheia, June 1984–June 1988). We also use the model to forecast the changes expected with treatment, which is to be continued through 1990, and the subsequent return towards pre-treatment conditions when the experiments end.

At Sogndal MAGIC satisfactorily simulates the observed concentrations of the three strong-acid anions (Cl^- , SO_4^{2-} , NO_3^-) (Fig. 1a). We assume chloride to be chemically inert and governed solely by atmospheric inputs and mixing with soil solution. In 1984 the chloride concentrations were high because of large inputs of sea salts; in the following years the chloride concentrations were markedly lower. This trend is correctly simulated by the model. In MAGIC sulphate is regulated by adsorption. The sulphate concentrations measured in runoff increase as the acid-addition experiment proceeds⁴. This trend is simulated by the model (Fig. 1a). The nitrate concentrations, both measured and simulated, were $<2 \mu\text{eq l}^{-1}$.

Concentrations of base cations are governed by the simultaneous action of several soil chemical processes. At Sogndal MAGIC correctly predicts the observed changes in concentrations of the base cations calcium, magnesium and potassium but fails to reproduce the large change in sodium concentration from the sea-salt-rich year 1984 to the years 1985–1987 (Fig. 1a). In the model, sodium concentrations are governed primarily by ion exchange, and the model indicates that this process should damp changes in sodium concentrations to a greater extent than we observe. The observed data shows that the sodium concentration follows that of chloride to a greater extent than expected if cation exchange were active. This indicates that the simple model with only one set of soil parameters is inadequate to describe the short-term changes. Other studies at Sogndal, including an experimental addition of dilute sea water, indicate that two types of soil water interact with precipitation to give runoff. Macropore water may contact soil with different cation exchange characteristics than the micropore water in the bulk soil⁸.

Alkalinity provides perhaps the best measure of success of the MAGIC predictions. We define alkalinity as the difference in sum of concentrations of base cations (Ca^{2+} , Mg^{2+} , Na^+ , K^+ , NH_4^+) and sum of strong acid anions (Cl^- , SO_4^{2-} , NO_3^-). Alkalinity is thus a calculated variable and involves all of the errors and uncertainties in these eight major ions. MAGIC correctly predicts the observed decrease in alkalinity from $\sim 10 \mu\text{eq l}^{-1}$ in 1983 to $\sim -5 \mu\text{eq l}^{-1}$ in 1987 (Fig. 1a). The four-year trends in pH and aluminium concentrations are also well described by MAGIC (Fig. 1a).

The model fits the observed 1983 data only when the weathering rates ($\text{meq m}^{-2} \text{yr}^{-1}$) for the four base cations are set within narrow bounds (Ca^{2+} , 3.4–4.1; Mg^{2+} , 0–0.3; Na^+ , 4.0–5.3; K^+ , 0–1.2; sum of the base cations, 7.4–10.5). These estimates of the weathering rates are close to those obtained by several other independent methods. These extremely low rates are reasonable for very thin organic-rich soils⁹.

Replication of the calibration routine produces narrow bands

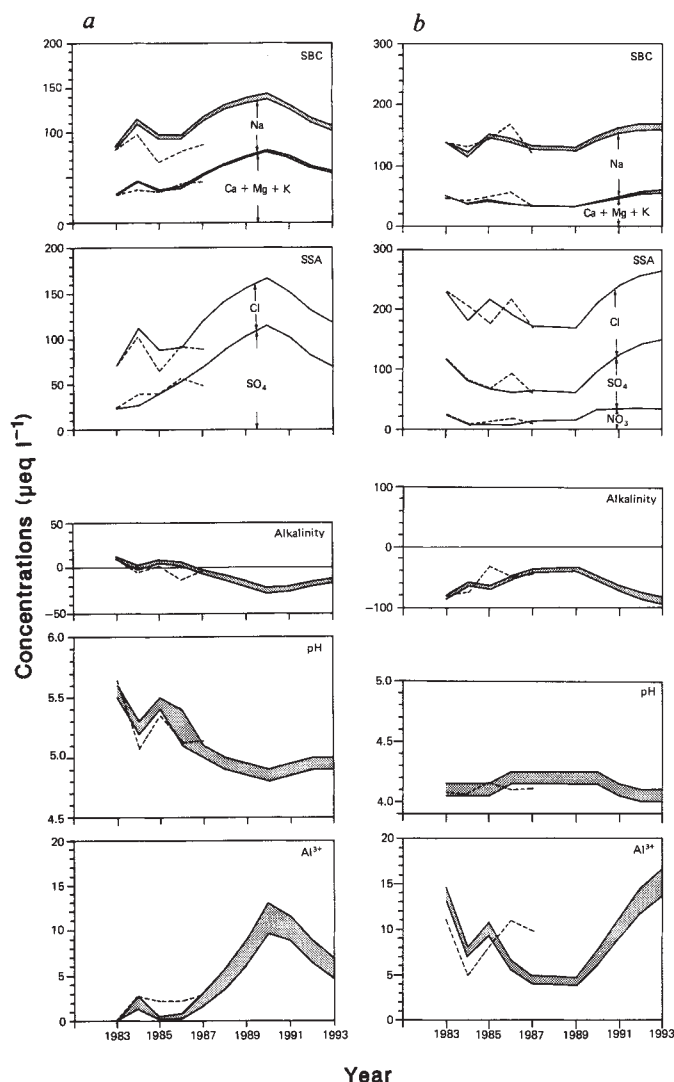


FIG. 1 a, Volume-weighted annual average concentrations of major constituents (SBC is the sum of basic cations and SSA is the sum of strong acids) in runoff from catchment SOG2 at Sogndal as measured (dashed lines) and simulated by the MAGIC model (solid lines). Bands of uncertainty (shaded) in the simulations are derived from eight replications of the optimization procedure. b, Volume-weighted annual average concentrations of major constituents in runoff from catchment KIM at Risdalsheia as measured (dashed lines) and simulated by the model (solid lines). Bands of uncertainty obtained as in a.

of uncertainty about the ten-year predictions (Fig. 1a). This indicates that our procedure for estimating weathering rate and original base saturation of the soil is robust. We also tested the sensitivity of the model to the uncertainties in the key soil parameters. Here we varied one parameter, recalibrated the model using the optimization routine and then compared the predictions produced by the model with the measured runoff chemistry. A 50% variation in the soil depth, p_{CO_2} , or concentrations of organic acids and an order-of-magnitude variation in aluminium solubility had no appreciable effect on the predictions. The predictions were not greatly affected by changes in the background levels of SO_4 and base cations in the precipitation. A change in any of these input parameters is compensated by change in one or more calibrated parameters in MAGIC, in particular the selectivity coefficients for the cation exchange reactions. The two major assumptions made in the MAGIC model—that the catchment is in a steady-state with respect to inputs and outputs before the onset of acid deposition (1843), and that the calculated soil chemistry and runoff chemistry match the measured chemistries in the calibration year (1983)—provide powerful constraints. The sensitivity analyses indicate that those sets of input and calibrated parameters for which these two constraints are met will produce similar predicted future trends in runoff chemistry.

For the acid-exclusion experiment at Risdalsheia, MAGIC again satisfactorily predicts the observed trends in concentrations of the strong-acid anions chloride, sulphate and nitrate (Fig. 1b). Here the major changes have been decreasing concentrations of sulphate and nitrate in runoff⁴. The model predictions for base-cation concentrations at Risdalsheia are better than those for Sogndal, especially with respect to sodium. This is probably because the year-to-year variation in sodium deposition at Risdalsheia has been considerably less than at Sogndal. The high chloride levels at Risdalsheia in 1987 were not due to seasalt inputs but rather to a technical malfunction of the ion-exchange system which resulted in a short-term input of HCl beneath the roof⁴.

The MAGIC model also simulates well the observed increase in alkalinity over the four-year experimental period (Fig. 1b). For both base cations and alkalinity the replications of the calibration procedure indicate relatively narrow bands of uncertainty related to estimation of weathering rates and initial base saturation. At Risdalsheia the range of weathering rates are Ca^{2+} : 3.9–7.4, Mg^{2+} : 0–2.9, Na^{+} : 2.0–5.6, K^{+} : 0–2.9, and sum of the base cations: 9.6–16.2 $meq\ m^{-2}\ yr^{-1}$.

Changes in both pH and aluminium concentration are also simulated well by MAGIC (Fig. 1b). Runoff at Risdalsheia contains high levels of dissolved organic matter which buffers pH. Organic acids are incorporated in MAGIC using the measured concentrations of total organic carbon with a charge density of 5 μeq per mg C and pK of 4.5 (Table 1)¹⁰.

We use the calibrated model to predict the effects to be expected with further years of treatment and the subsequent return to pre-treatment conditions at both Sogndal and Risdalsheia. At Sogndal the model indicates that three more years of acid addition will result in a further doubling of sulphate concentrations in the runoff and consequently a decrease in alkalinity to $-20\ \mu eq\ l^{-1}$ and pH to 4.9. The aluminium concentration will increase dramatically to $\sim 11\ meq\ l^{-1}$ (Fig. 1a). The recovery following cessation of treatment in 1991 is predicted to be slower than the rate of acidification. This hysteresis is due to the fact that the depletion of the pool of exchangeable base cations in the soil (soil acidification) is closely related to the flux of sulphate through the soils, but the rebuilding of this pool is accomplished primarily by the release of base cations through weathering. The weathering rate at Sogndal is only $\sim 10\ meq\ m^{-2}\ yr^{-1}$, appreciably lower than the 20–60 $meq\ m^{-2}\ yr^{-1}$ increased flux of base cations accompanying the sulphate.

At Risdalsheia the model indicates that two more years of

TABLE 1 Catchment and soil parameters

Parameter	Sogndal	Risdalsheia
Soil depth (m)	0.30	0.11
Porosity (%)	0.5	0.5
Bulk density ($kg\ m^{-3}$)	618	400
Cation exchange capacity ($meq\ kg^{-1}$)	37	82
SO_4 adsorption half saturation ($meq\ m^{-3}$)	60	60
SO_4 adsorption maximum capacity ($meq\ kg^{-1}$)	6.0	6.0
Solubility $Al(OH)_3$	8.8	7.6
Selectivity coefficient Al–Ca	–1.46	–1.87
Selectivity coefficient Al–Mg	–0.35	–0.17
Selectivity coefficient Al–Na	–1.77	–2.18
Selectivity coefficient Al–K	–6.12	–6.33
Total organics, solution ($mmol\ m^{-3}$)	20	75
Organic pK_1	4.0	4.5
Organic pK_2	8.0	8.0
CO_2 , soil air (atm)	0.006	0.0015
Soil temperature ($^{\circ}C$)	6.6	7.0
CO_2 , stream water (atm)	0.0011	0.0008
Stream temperature ($^{\circ}C$)	6.6	7.0
Ca saturation, soil 1983/1984 (%)	17.5	5.5
Mg saturation, soil 1983/1984 (%)	3.9	2.7
Na saturation, soil 1983/1984 (%)	2.0	1.4
K saturation, soil 1983/1984 (%)	3.0	1.9
Total base saturation 1983/1984 (%)	26.4	11.5

acid exclusion will cause a significant decrease in aluminium concentration. The pH and alkalinity will increase only slightly because of the buffering by organic acids (Fig. 1b). After treatment ends, the return to pre-treatment conditions will be considerably more rapid than at Sogndal, because of the much thinner soils and lower base saturation.

Application of the MAGIC model to the whole-catchment experimental manipulations of the RAIN project shows that the response of water and soil acidification to large and rapid changes in acid deposition can be predicted. The aggregation approach used in the MAGIC model to describe a complex system of spatially varying chemical and biological reactions produces predicted trends in runoff chemistry that agree satisfactorily with observed trends. The model is not designed to deal with short-term fluctuations in runoff chemistry that are due to changing hydrological pathways, seasonal variations in precipitation chemistry and other such cyclic events. MAGIC focuses on the year-to-decade changes in runoff chemistry, and the gross trends in the four-year data from the RAIN project are indeed reflected in the model predictions. These results reinforce other evaluations of the model obtained by comparison with palaeolimnological reconstructions of lake acidification^{11,12} and changes in regional lake chemistry¹³. Together these applications give us confidence in the use of the MAGIC model as a robust tool for predicting future soil and water acidification following changes in acid deposition. □

Received 22 June; accepted 13 November 1989.

1. Cosby, B. J. *et al.* *Wat. Resour. Res.* **21**, 51–63 (1985).
2. Cosby, B. J. *et al.* *Wat. Resour. Res.* **21**, 1591–1601 (1985).
3. Wright, R. F. & Cosby, B. J. *Atmos. Environ.* **21**, 727–730 (1987).
4. Wright, R. F. *et al.* *Nature* **334**, 670–675 (1988).
5. Lotse, E. & Ottaabong, E. *Acid Rain Res. Rep.* 8/85 (Norwegian Institute for Water Research, Oslo, 1985).
6. Lotse, E. *Acid Rain Res. Rep.* 18/89 (Norwegian Institute for Water Research, Oslo, 1989).
7. Bettelheim, J. & Littler, A. *CEGB Rep. PL-GS/E/1/79* (Central Elect. Generating Board, London, 1979).
8. Wright, R. F. *et al.* *Nature* **334**, 422–424 (1988).
9. Frogner, T. *Geochim. cosmochim. Acta* (in press).
10. Wright, R. F. *Wat. Air Soil Pollut.* (in press).
11. Wright, R. F. *et al.* *Wat. Air Soil Pollut.* **30**, 367–380 (1986).
12. Neal, C. *et al.* *Nature* **334**, 109–110 (1988).
13. Hornberger, G. M., Cosby, B. J. & Wright, R. F. *Wat. Resour. Res.* **25**, 2009–2018 (1989).

ACKNOWLEDGEMENTS. This work was supported by the Norwegian Institute for Water Research, the Royal Norwegian Council for Scientific and Industrial Research, the Surface Water Acidification Programme (SWAP) (The Royal Society, the Norwegian Academy of Science and Letters, and the Royal Swedish Academy of Sciences) and the NATO Science Committee.

Palaeoclimatological and chronological implications of the Vostok core dust record

J. R. Petit^{*†}, L. Mounier^{*}, J. Jouzel^{†*},
Y. S. Korotkevich[‡], V. I. Kotlyakov[§] & C. Lorius^{*}

^{*}Laboratoire de Glaciologie et Géophysique de l'Environnement, BP 96, 38402, St Martin d'Heres Cedex, France

[†]Laboratoire de Géochimie Isotopique, IRDI/DLPC CEN Saclay, 91191 Gif sur Yvette, France

[‡]Arctic and Antarctic Research Institute, Beringa Street 38, 199226 Leningrad, USSR

[§]Institute of Geography, Academy of Sciences of the USSR, Staromonetny per. 29, Moscow 109017, USSR

THE 2,083-m Vostok ice core recovered by the Soviet Antarctic expeditions has provided much information of climatic and environmental interest, for a period covering a full glacial-interglacial cycle¹⁻⁶. Here we present and discuss the dust record obtained down to 2,202 m, the final depth to which this core was extended in 1986. First, we document the fact that major changes in aeolian deposits, as recorded in the Vostok core, appear to be of global significance and confirm the existence of a link between high-latitude aeolian deposits and the Earth's orbital parameters^{7,8}. Second, we propose atmospheric dust as a stratigraphic marker to compare timing with other records of palaeoclimate, and use the magnetic-susceptibility profile measured along the RC11-120 Indian Ocean core⁹ for this purpose, assuming that major dust events correspond to common aeolian inputs. This approach indicates that the Vostok and marine records were roughly in phase at the previous glacial-interglacial transition.

We sampled the Vostok core every 8 m and performed microparticle studies on ice sections 15–20 cm long, following a procedure based on optic and Coulter countings¹⁰. The upper part of the core (above 1,000 m) shows many cracks that prevent the use of the samples for Coulter counting because they may have been contaminated by drilling fluid. We then took advantage of the excellent correlation ($r=0.90$ on 60 crack-free samples) between the aluminium content for which concentration is not perturbed by drilling fluid and the microparticle concentration^{2,11} to estimate the microparticle concentration in the upper 1,000 m of the core, assuming that the aluminium content of the dust is 7%. This combination of methods allowed us to produce a reliable and relatively detailed dust record from the surface down to 2,202 m (276 values). Figure 1a shows the dust volume concentration as a function of depth as well as the Vostok isotope stages as defined in ref. 1.

We used a timescale established using an ice-flow model¹, to date the core to a depth of 2,202 m, which corresponds to an age of ~185 kyr. We used deuterium measurements, which have also been recently extended down to 2,202 m (1 m of ice analysed for every 2 m, so far), to estimate the surface temperatures at the Vostok site³ (Fig. 2a). The newly analysed part of the record (below 2,083 m) corresponds to the penultimate glacial period with a morphology fairly similar to the latter part of the last glacial period.

We express the dust data in terms of the annual flux calculated by multiplying the volume concentration and the annual snow accumulation which we estimated from the isotopically derived temperature profile^{1,12}. Figure 2b shows a smoothed curve of the dust flux with respect to time. The lowest concentrations and fluxes of dust occur in the interglacial periods (Holocene or stage A and G, ~130 kyr BP). Large dust spikes occur between 330 and 550 m (stage B, ~20 kyr BP), 860 and 1,000 m (mostly in stage D, ~60 kyr BP) and below 1,950 m (stage H, ~140–180 kyr BP) a pattern similar to that previously described for the aluminium and non-sea-salt calcium (excess calcium)

contents^{2,5}). Compared with the Holocene value, the dust flux is greater respectively by a factor of 15, 11 and 9. The absence of a peak during the cold stage F (~110 kyr BP), previously noted for other terrestrial input indicators^{2,5}, is confirmed by this dust record. We also note that for each glacial-interglacial transition (H/G, B/A) the dust change (from high to low fluxes) clearly leads the temperature change with a return to low dust fluxes at mid-transition (Fig. 2a, b).

Microparticles found in Antarctic ice represent the small-size fraction ($r \approx 1 \mu\text{m}$) of continental dust comprising various aluminosilicates including mainly clay minerals, quartz and some feldspars^{10,13-16}. There is also a volcanic-aerosol contribution, but this never exceeds 20%, and we have not taken it into account in our general interpretation of the dust record¹⁰. The semi-arid and arid areas of the three continents of the Southern Hemisphere, the non-glaciated areas of Antarctica and the continental shelves—which were progressively exposed to wind erosion owing to sea-level lowering—are the potential sources of the dust deposited in Antarctica¹⁰. The interpretation that the Antarctic dust deposition depends on aridity, wind speed over source areas and meridional transport efficiency^{16,17} may be extended for the entire Vostok dust record. The absence of a dust peak during stage F (~110 kyr BP) may therefore be due either to insufficient wind speeds over the source areas or to the fact that, unlike the case for cold stages B, D and F, the sea level was still high and the aridity was then not very different from that of interglacial periods. Indeed, in western Europe the climate was very humid and markedly cold during the final part of the Eemian¹⁸.

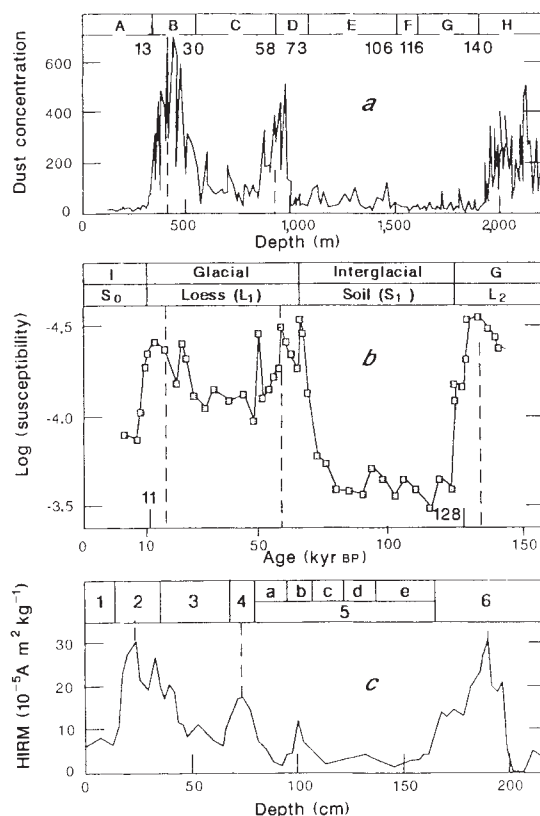


FIG. 1 Comparison of dust records from different locations over the last climatic cycle. *a*, Vostok dust concentration expressed in volume ($10^{-9} \text{ cm}^3 \text{ g}^{-1}$) plotted against depth. Isotope stage and age are from ref. 3. *b*, Magnetic susceptibility measured on Xifeng loess sequence plotted against age (from ref. 7). (Stratigraphy: L = loess, S = Soil. 11 kyr and 128 kyr are the ages assigned to soil/loess boundary.) These two fixed points and that corresponding to the Brunhes/Matuyama boundary (730 kyr BP) were used to establish²⁷ the timescale. *c*, Magnetic remanence parameter (HIRM) on S8-78 ($41^{\circ}43' \text{N}$, $20^{\circ}23' \text{W}$) in North Atlantic against depth with marine stages (from ref. 28).

For spectral analysis of the dust record we used a multitaper method^{19,20} that provides an estimation of the variance and a statistical test defining the reliability of each spectral peak. The spectra show that most of the variance is near the obliquity and precession frequency bands (as also found using both the maximum entropy method¹⁰ and the calcium profile⁷), but a 90% confidence level is reached only for the precession, which shows two significant peaks (21.6 and 18.5 kyr), in fair agreement with the double precession peak (19 and 23 kyr) derived from insolation curves²¹. These spectral analyses indicate that dust deposition in polar regions is through changes in aridity and/or wind intensity (or in both), influenced by the orbital forcing. A recent study of loess deposition in Alaska⁸ independently arrived at the same conclusion.

The Vostok ice-core data for the last glacial maximum compare well with those from other deep ice cores. Both the Antarctic and Greenland records are characterized by high dust content^{16,22,23}. Comparison for the earlier period is difficult because other cores are either too short (Dome C) or not dated with a sufficient accuracy over the glacial period (Camp Century, Byrd, Dye 3). Comparison with Southern Hemisphere continental dust records is restricted because data for aeolian dust transport and continental aridity are generally limited to the last ~40 kyr^{24,25}, except for a few series (J. Wasson, unpublished data) that are discontinuous and only qualitative. We then focus the comparison on long and continuous series from Northern Hemisphere coming either from a continental loess record in China²⁶ or from a marine record from the North Atlantic²⁷.

The Chinese loess plateau is a deposit, over 200 m thick, in which layers of silt transported by wind from the Central and Gobi deserts alternate with ancient soils for the last 2.5 million years²⁶. Loess was deposited in cold arid and windy glacial climates whereas the embedded soils were formed under mild and wet climates. Thus the alternating sequence of loess and soil layers records oscillations between relatively dry and more humid conditions^{26,28}. The magnetic susceptibility was found to be higher in soils and lower in loess and can be taken as an indicator of the dust source strength in central Asia. During the past 180 kyr of the Xifeng series (Fig. 1b), there were three periods of aridity, the pattern of which is broadly comparable with that in the Vostok core (Fig. 1a). From ~110 to 80 kyr BP, the stratigraphy of the Chinese core shows a soil layer indicative of a wetter climate without dust production, which is also in agreement with the Vostok record.

In the North Atlantic record, information²⁷ about continental dust was derived from both mineralogical and magnetic studies. The magnetic remanence (HIRM) curve²⁷ (Fig. 1c) is an index of the aeolian contribution of the northern part of Africa to the sedimentation in this area. High aeolian contributions for glacial marine stages 2, 4 and 6 are confirmed and again there is a general correspondence between this HIRM curve and the Vostok dust record, except for an event occurring during marine stage 5b.

Establishing a stratigraphic correlation between the Vostok core and one of the well documented Northern Hemisphere record is tempting, but it is not yet clear that these results are representative on a global scale. Taking advantage of the relationship between magnetic susceptibility and aeolian deposition for loess deposits²⁶ and marine sediments²⁷, we examined how the magnetic-susceptibility record in the Indian Ocean core RC11-120 (79°52 E, 43°21 S) compares with the Vostok dust record. Magnetic-susceptibility measurements in this core have been performed and discussed⁹. In Fig. 2c, we express this parameter as an annual flux by multiplying the susceptibility data and the annual sedimentation rate, as deduced from the depth-age relationship²⁹. This curve, with three well marked peaks during marine stages 2, 4 and 6 (defined as in the $\delta^{18}\text{O}$ SPECMAP (Spectral Mapping) record reproduced in Fig. 2d) agrees well with the Vostok dust flux (Fig. 2b) which, from the above discussion, may reflect common aeolian input. Although

the contribution of river input or of ice rafting from icebergs may be neglected³⁰, different interpretations are possible, such as the northward transport, by the Antarctic bottom-water current³⁰, of smectite (a mineral generally associated with ferromagnetic minerals) from volcanic areas such as the Kerguelen plateau. Higher magnetic susceptibility during stages 2, 4 and 6 could thus be due to higher rate of deep-water formation, resulting in greater erosion of the ocean floor and in more efficient transport of the eroded material by the Antarctic bottom-water circulation. At present there is insufficient information on deep-water circulation changes in the Indian Ocean either to confirm or to discount such a factor.

There are, however, several indications¹⁰ of an aeolian contribution at this site. One such indication (Fig. 2c) comes from quartz measurements³¹ on samples from an area that includes the site at which RC11-120 was taken. This would¹⁰ correspond to a quartz flux about four times higher during the last glacial maximum than during the Holocene. Although we must be cautious in interpreting the RC11-120 magnetic-susceptibility record as a dust record, we suggest that the major susceptibility peaks in RC11-120 are partly due to continental dust inputs and that these well marked events (marine stages 2, 4 and 6) are of sufficiently large geographical extent to be recorded both in RC11-120 and in Vostok. Although further analysis to examine, for example, chemical composition relationships would be very useful, similarities between the magnetic susceptibility (Fig. 2c) and the dust (Fig. 2b) records indirectly support this hypothesis.

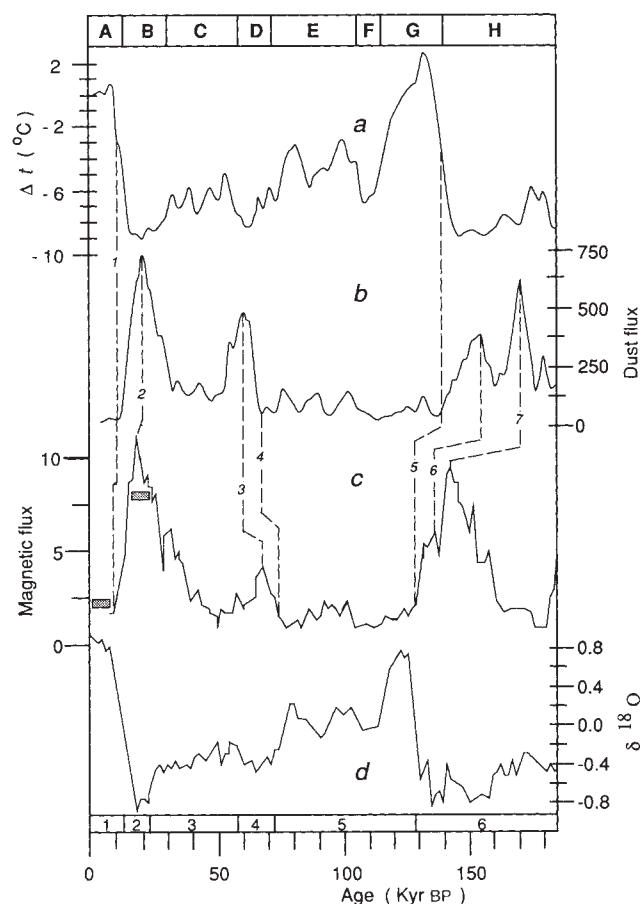


FIG. 2 Comparison between the Vostok and RC11-120 records over the last 180 kyr and the possible correspondence through the dust events. *a*, Vostok temperature as derived from stable-isotope measurements. *b*, Vostok dust flux ($\times 10^{-9} \text{ cm yr}^{-1}$), after smoothing by a cubic spline function. *c*, RC11-120 magnetic-susceptibility 'flux' and the estimated quartz flux variation for the Holocene and the last glacial maximum (in relative units). *d*, SPECMAP normalized oxygen isotope curve (from ref. 29). The isotope stages for Vostok (from ref. 1) and the marine stages are also shown.

We therefore accept, provisionally, that these major dust events provide a common stratigraphic marker which allows the direct correlation of the marine and Vostok records. For this purpose, we selected major events recognizable in both records (Fig. 2) that correspond either to maximum values (events 2, 3, 6 and 7) or to the beginning (event 4) or termination (events 1 and 5) of a peak.

To test the validity and limits of this approach we first examine the last climatic transition. Low dust contents at Vostok (Fig. 2b) occur at mid-transition (12.2 kyr BP) whereas at RC11-120 (Fig. 2c) the minimum in the magnetic susceptibility record occurs about 3.5 kyr later. So large a difference for such a recent event may be explained only partially by dating inaccuracy, either on Vostok or RC11-120; hence we consider this value as a limit of the accuracy of our stratigraphic approach.

Discrepancies between the Vostok core and marine core chronologies before 110 kyr BP have already been noted and discussed³. They are due to the different duration of the last interglacial period, which was about twice as long in the Vostok chronology (stage G) as it is in the marine chronology. As a result, at mid-transition from glacial to interglacial (H/G or 6/5), the Vostok temperature leads the global ice volume (as given by the $\delta^{18}\text{O}$ SPECMAP record) and the surface temperature at RC11-120 by respectively 10 and 7 kyr (ref. 3). For this transition the time difference between the end of the dust peak (Fig. 2b) and the end of the magnetic 'flux' peak is ~ 10 kyr and using this level as a time marker would put the Vostok temperature record practically in phase with the marine $\delta^{18}\text{O}$ record (Fig. 2). The Vostok temperature would then lag the RC11-120 sea surface temperature by ~ 3 kyr, but this value is not significant given the limit of accuracy mentioned above. Within this limit, we conclude that there was no significant phase shift between the Vostok temperature and the marine records during the previous glacial-interglacial transition. We also note that the time interval between levels 4 and 5 is 30% less in the Vostok than in the marine chronology (from 138 to 66 kyr BP and from 128 to 73 kyr BP respectively). This might reduce the discrepancy between the duration of the last interglacial in the two records, but we note that the 'dust' event during marine stage 4 is less well characterized than those of stages 2 or 6.

Our approach is independent of absolute dating. We note, however, that recent results obtained on coral-reef terraces place the peak of the last interglacial around 126 kyr BP (E. Bard, B. Hamelin, Y. Lao, R. F. Anderson and R. G. Fairbanks, unpublished results) confirming the deep-sea-core dating obtained by orbital tuning²⁹. The Vostok timescale would thus be in error. This would be a strong argument against the evidence from a continental sequence³² for a much earlier interglacial period.

Our stratigraphic approach relies on the critical assumption that magnetic susceptibility in RC11-120 reflects the aeolian input. Our results must be confirmed and all the information relevant to the Vostok and marine chronologies combined before a revised Vostok timescale may be established. We intend to make further magnetic measurements and mineralogical investigations to ascertain the origin of the magnetic material and to extend the technique of using dust as a stratigraphic marker. □

12. Lorius, C., Raisbeck, G., Jouzel, J. & Raynaud, D. in *The Environmental Record in Glaciers and Ice Sheets* (eds Oeschger, H. & Langway, C. C. Jr) 343–361 (Wiley, Chichester, 1989).
13. Briat, M., Royer, A., Petit, J. R. & Lorius, C. *Ann. Glaciol.* **3**, 27–31 (1982).
14. Gaudichet, A., Petit, J. R., Lefevre, R. & Lorius, C. *Tellus* **B38**, 250–261 (1986).
15. Gaudichet, A. et al. *Geophys. Res. Lett.* **15**, 1471–1474 (1988).
16. Petit, J. R., Briat, M. & Royer, A. *Nature* **293**, 391–394 (1981).
17. Lorius, C., Raynaud, D., Petit, J. R., Jouzel, J. & Merlivat, L. *Ann. Glaciol.* **5**, 88–95 (1984).
18. Guiot, J., Pons, A., De Beaulieu, J. L. & Reille, M. *Nature* **338**, 309–313 (1989).
19. Thomson, D. J. *Proc. IEEE* **70**, 1055–1096 (1982).
20. Yiou, F. et al. in *Interaction of the Global Carbon Cycle and Climate System, Electric Power Res. Inst. Rep.* (ed. Keir P.) (Kluwer, in the press).
21. Berger, A. I. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
22. Thompson, L. G. & Mosley-Thompson, E. *Science* **212**, 812–815 (1981).
23. Hammer, C. et al. *Geophys. Monogr.* **33**, 90–94 (1985).
24. Heusser, C. J. *Quat. Res.* **16**, 293–321 (1981).
25. Bowler, J. M. *Earth Sci. Rev.* **12**, 279–310 (1976).
26. Kukla, G. *Quat. Sci. Rev.* **6**, 191–219 (1987).
27. Robinson, S. G. *Phys. Earth planet. Inter.* **42**, 22–47 (1986).
28. Hovan, S. A., Rea, D. K., Pisias, N. G. & Shackleton, N. J. *Nature* **340**, 296–298 (1989).
29. Martinson, D. G. et al. *Quat. Res.* **27**, 477–482 (1987).
30. Kolla, V., Henderson, L. & Biscaye, P. *Deep Sea Res.* **23**, 949–961 (1976).
31. Kolla, V. & Biscaye, P. *J. sedim. Pedology*, **47**, 642–649 (1977).
32. Winograd, I. J., Barney, J., Zabo, B. J., Coplin, T. B. & Riggs, A. C., *Science*, **242**, 1275–1280 (1988).

ACKNOWLEDGEMENTS. We thank M. Cresceur, P. Yiou and C. Genthon for the implementation of the MTM, C. Ritz for the age-depth Vostok model, and M. Ghil, F. Grousset, J. Imbrie, D. Kent, D. Raynaud and W. Ruddiman for discussions.

The role of strain energy in creep graphitization of anthracite

J. V. Ross & R. M. Bustin

Department of Geological Sciences, The University of British Columbia, Vancouver, Canada, V6T 2B4

RECENT research on ceramics and natural minerals has demonstrated that non-hydrostatic stress can affect some polymorphic transitions and can increase reaction rates^{1,2}. One such example is the graphitization of anthracite. Under natural conditions graphite forms at temperatures of 300–500 °C and confining pressures of ~ 500 MPa (refs 3–9). But in simple heating experiments at ambient pressure and high confining pressure (up to 1 GPa), temperatures of $\sim 2,000$ °C are required for graphite formation^{10–13}. Here we report creep experiments on natural anthracite at temperatures of 300–600 °C, using variable strains and strain rates and a constant confining pressure of 500 MPa. The experiments yield an apparent activation energy of 68.6 kJ mol^{-1} for the steady-state process(es) leading to graphite formation. This value is in marked contrast with simple heating experiments, which require an activation energy of $1,000 \text{ kJ mol}^{-1}$ (ref. 10). We suggest that in our experiments, and also under natural conditions, graphitization is facilitated by available strain energy associated with non-hydrostatic stress; such stresses typify conditions of natural graphitization.

Naturally occurring graphite has been documented from many localities (usually under conditions of synkinematic metamorphism associated with high strain) ranging from upper greenschist facies to lower amphibolite facies and above^{3–7}. To test the hypothesis that the strain-energy component of natural deformation is a very important part of the energy required for natural graphitization we planned a series of axial-symmetric deformation experiments on natural samples of vacuum-dried anthracite, which were maintained at 110 °C for a minimum of 1 week. These experiments were conducted under reducing conditions, comparable to deformation and metamorphism under natural conditions: methane was evolved during the experiments. We used two types of confining-pressure medium—a solid medium (NaCl) for high-strain experiments ($> 15\%$ axial shortening)¹⁴ and argon gas for low-strain experiments¹⁵. All tests were conducted in compression using a confining pressure of 500 MPa, a temperature range of 350–600 °C and constant strain rates of

Received 29 June; accepted 8 November 1989.

1. Lorius, C. et al. *Nature* **316**, 591–596 (1985).
2. De Angelis, M., Barkov, N. I. & Petrov, V. N. *Nature* **325**, 318–321 (1987).
3. Jouzel, J. et al. *Nature* **329**, 403–408 (1987).
4. Barnola, J. M., Raynaud, D., Korotkevitch, Y. S. & Lorius, C. *Nature* **329**, 408–414 (1987).
5. Legrand, M., Lorius, C., Barkov, N. I. & Petrov, V. N. *Atmos. Envir.* **22**, 317–331 (1988).
6. Raynaud, D., Chappellaz, J., Barnola, J. M., Korotkevitch, Y. S. & Lorius, C. *Nature* **333**, 655–657 (1988).
7. Legrand, M., Delmas, R. J. & Charlson, R. J. *Nature* **334**, 418–420 (1988).
8. Beget, J. E. & Hawkins, O. B. *Nature* **337**, 151–153 (1989).
9. Kent, D. V. *Nature* **299**, 538–539 (1982).
10. Mounier, L. thesis, Université Scientifique Technologique et Médicale de Grenoble (1988).
11. De Angelis, M. et al. *Atmos. Chem.* **1**, 215–239 (1984).

$1.3 \times 10^{-6} \text{ s}^{-1}$ to $1.3 \times 10^{-4} \text{ s}^{-1}$. After being corrected for apparatus and jacket effects^{16,18}, all stress-strain curves have similar geometry, with the stress difference at any strain decreasing with increasing temperature or with decreasing strain rate; steady-state conditions were achieved at $\sim 5\%$ axial strain. Values of stress difference, measured at 10% axial strain, for different strain rates are shown in Table 1 and are plotted as log (stress difference) against $-\log$ (strain rate) in Fig. 1. We use a linear least-squares minimization procedure¹⁶ to fit the data to a power-law creep equation of the form¹⁵

$$\ln \dot{\epsilon} = \ln A - Q/RT + n \ln \sigma$$

where A is the pre-exponential factor, Q is the apparent activation energy, n is the stress exponent, $\dot{\epsilon}$ is the strain rate, T is the temperature (K) and σ is the stress. The best-fit (correlation coefficient, $r = 0.92$) isotherms through the data points have a slope n and increase their spacing slightly towards the lower temperatures. The parameters of this fit (confining pressure of 500 MPa) are $A = 10^{0.99 \pm 0.15} \text{ s}^{-1} \text{ MPa}^{-n} \text{ K}$, $Q = 68.6 \pm 1.2 \text{ kJ mol}^{-1}$ and $n = 3.3 \pm 0.2$, where the errors quoted are ± 2 standard deviations (95% confidence limits).

We have examined all of the deformed samples and undeformed specimens using reflected light. As previously discussed¹⁹, the maximum reflectance increases with increasing strain and at about 23% strain is preferentially developed parallel to the plane of flattening (plane perpendicular to the maximum compressive stress, or load axis); and the bireflectance (the difference between maximum and minimum reflectance) increases.

Reordering of the molecular structure accompanying deformation suggested by optical property(s) is also evident in transmission electron microscopy (TEM) of the undeformed and deformed samples. Undeformed samples are characterized by elongated (10–20-nm) pores lying in the bedding plane and by pore walls composed of 'stacks' of basic structural units (BSUs) or aromatic layers that are 1–5 nm thick and up to 10 nm long (Fig. 2). These configurations are similar to those previously described^{6,13}. Using TEM we examined two samples that were shortened by $>25\%$ at 600 °C using a strain rate of 10^{-5} s^{-1} , and found them to be heterogeneous in structure. The pore size (5–10 nm in length) and density were reduced and the BSUs were preferentially oriented in the plane of flattening. The BSUs of which the pore walls were composed, although turbostratic³, had a more planar or linear nature in the flattening plane.

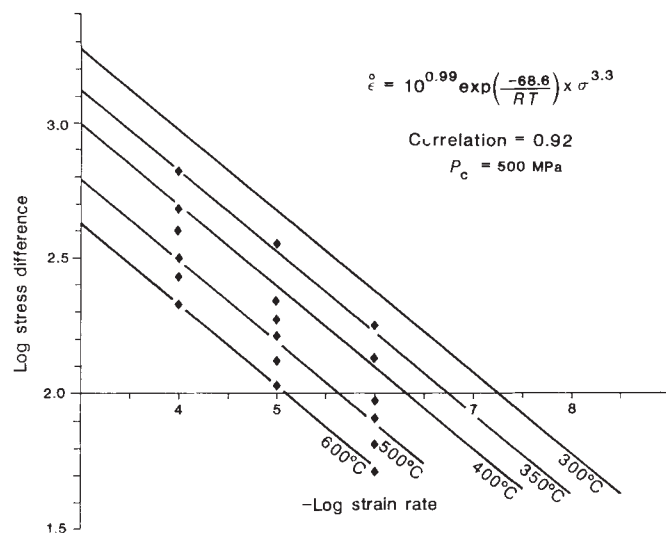


FIG. 1 The logarithm of stress difference plotted against the logarithm of the strain rate for steady-state data from constant-strain-rate, constant-temperature experiments (data shown in Table 1). All experiments at 500-MPa confining pressure and stress difference measured at 10% strain.

TABLE 1 Stress difference at various strain rates

Temperature °C	10^{-4} s^{-1}	Strain rate 10^{-5} s^{-1}	10^{-6} s^{-1}
350	660	354	177
400	478	220	135
450	398	186	95
500	316	162	81
550	269	131	66
600	213	107	52

Stress difference ($\sigma_3 - \sigma_1$) (in MPa) measured at 10% axial strain in experiments at a constant confining pressure of 500 MPa.

Selected-area diffraction (SAD) imaging in the dark field shows many microstructural orientations within the specimen (Fig. 3a). Other areas show strong preferred orientation in SAD images of their (002) lattice fringes (Fig. 3b), indicating that a high degree of anisotropy is associated with strains of over 23%, similar to the structure of graphite produced at high temperature and ambient pressure by simple heating. At some positions, where the anisotropy seems to be greatest, dark-field imaging shows a crystalline SAD pattern (Fig. 3c) and the bright-field images are characterized by Bragg and Moiré fringes (Fig. 3d), indicative of graphite lamellae that may be either single crystals or overlapping crystallites.

TEM of undeformed and deformed specimens gives insight into the process(es) operating during steady-state deformation: that results in the formation of graphite. Pores have coalesced, rotated and become aligned in the flattening plane of the experiments. Furthermore, the BSUs or stacks of aromatic layers have become more ordered, with the eventual production of graphite. This long-range ordering of the structure(s) is associated with large values of reflectivity and bireflectance of highly strained samples. The graphite produced in the deformation experiments is comparable to that produced by simple heating under ambient pressure. We suggest that the activation energy of $1,000 \text{ kJ mol}^{-1}$ associated with simple heating experiments is the sum of the thermal activation energy 68.6 kJ mol^{-1} for diffusion-biased creep¹⁷ and the strain energy imparted to the sample(s) during a minimum of 23% axial shortening.

Recent reviews²⁰ show that most solid-state transformations involve nucleation and growth and, with a suitable driving stress,

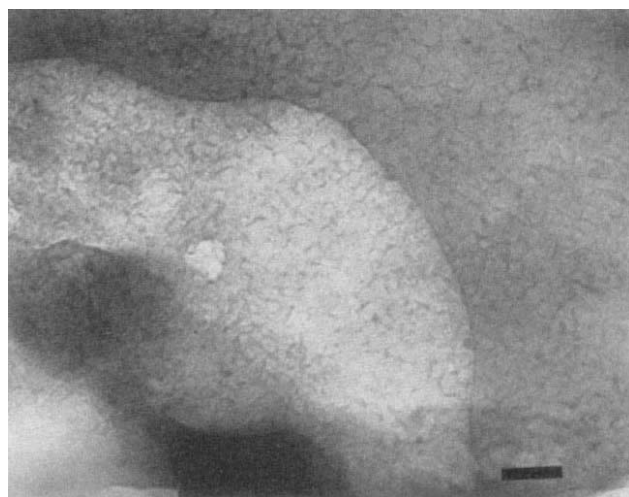


FIG. 2 Transmission electron micrograph (dark field) of the starting material (undeformed specimen of anthracite), showing the network of (002) images indicating the microporosity and sections of the aromatic layers or BSUs of which the pore walls are composed. Scale bar is 1,000 nm.

can occur at low temperatures requiring little thermal activation. These transformations therefore occur by the well-known effect of shear stress on reaction kinetics observed in many ceramics and minerals²¹⁻²³. In natural deformation the production of graphite at the much lower temperatures of 400–500 °C is probably aided by shear-strain effects accompanying natural deformation. In nature, shear strains are often in excess of $\gamma = 2$ ($\gamma = 2$ corresponds to axial shortening of ~60%). This process

of graphite formation is also probably related to the diffusional release of H^+ and CH_4 molecules associated with structural rearrangement and alignment of BSUs. The rest of the activation energy needed for the graphitization process comes from shear strain, which facilitates the straightening of the aromatic layers to a more uniform graphite packing and, on the scale of the optical microscope, also leads to an increase in reflectivity and birefractance. □

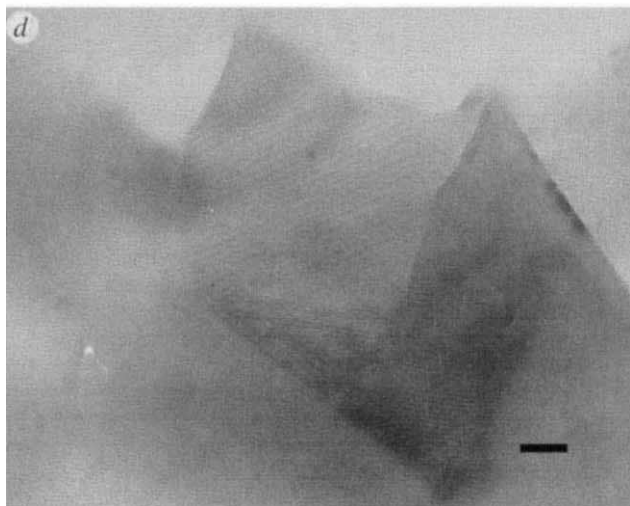
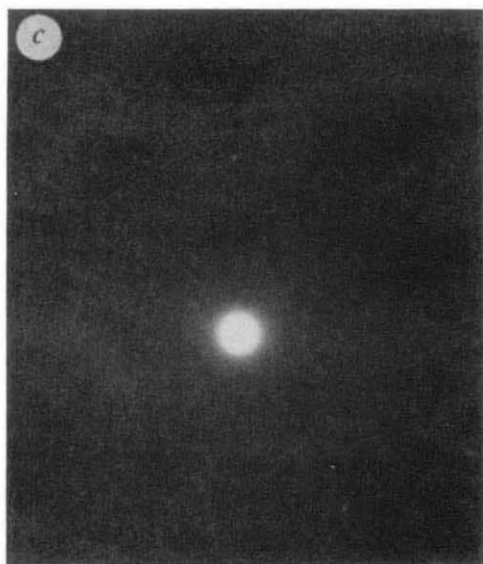
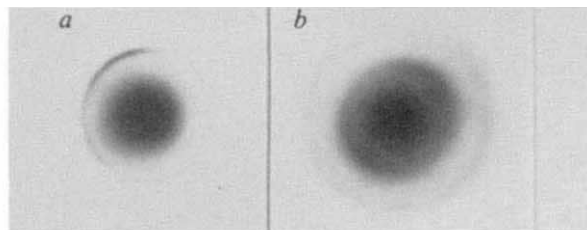


FIG. 3 *a*, Selected-area diffraction pattern (SAD—dark field) of highly deformed specimen, showing (002) ring containing 'spotty' reflections, indicating that the layers are turbostratic. *b*, SAD (dark-field) pattern from same specimen showing well developed arcuate pattern of the (002) ring, indicating a high degree of preferred orientation of the aromatic layers. *c*, Dark-field SAD pattern of crystalline graphite. *d*, Bright-field electron micrograph of region in *c*, showing graphite lamellae having 0.335-nm spacing, with Bragg fringes and poorly developed Moiré fringes, resulting from partial overlapping and rotation of graphite lamellae. Scale bar is 10 nm.

Received 31 July; accepted 10 November 1989.

1. Poirier, J.-P. *Creep of crystals* (Cambridge University Press, London, 1985).
2. Snow, E. & Yund, R. A. *J. metamorphic Geol.* **5**, 141–153 (1987).
3. Franklin, R. E. *Proc. R. Soc.* **209**, 196–218 (1951).
4. Evans, E. L., Jenkins, J. L. & Thomas, J. M. *Carbon* **10**, 637–642 (1972).
5. Diessel, C. F. K., Brothers, R. N. & Black, P. M. *Contr. Miner. Petrol.* **68**, 63–78 (1978).
6. Deurbergue, A., Oberlin, A., Oh, J.-H. & Rouzaud, J. N. *Int. J. Coal Geol.* **8**, 375–393 (1987).
7. Oh, J.-H. thesis, Univ. d'Orléans (1987).
8. Landis, C. A. *Contr. Miner. Petrol.* **30**, 34–45 (1971).
9. Okuyama-Kusunose, Y. & Itaya, T. *J. metamorphic Geol.* **5**, 121–139 (1987).
10. Bonijoly, M., Oberlin, M. & Oberlin, A. *Int. J. Coal Geol.* **1**, 283–313 (1982).
11. Noda, T. & Kato, H. *Carbon* **3**, 289–297 (1965).
12. Oberlin, A. *Carbon* **22**, 521–541 (1984).
13. Rouzaud, J. N. thesis, Univ. Orléans (1984).
14. Green, H. W., Griggs, D. T. & Christie, J. M. in *Experimental and Natural Rock Deformation* (ed. Paulitsch, P.) 272–335 (Springer, New York, 1970).
15. Heard, H. C. *Geol. Soc. Am. Mem.* **79**, 193–244 (1960).
16. Heard, H. C. & Raleigh, C. B. *Geol. Soc. Am. Bull.* **83**, 935–956 (1972).
17. Carter, N. L. *Rev. Geophys. Space Phys.* **14**, 301–360 (1976).
18. Parrish, D. K., Krivz, A. & Carter, N. L. *Tectonophysics* **32**, 183–207 (1976).
19. Bustin, R. M., Ross, J. V. & Moffat, I. *Int. J. Coal Geol.* **6**, 343–351 (1986).
20. Rubie, D. C. & Thompson, A. B. *Adv. phys. Geochem.* **4**, 27–79 (1985).
21. Poirier, J.-P. *J. geophys. Res.* **87**, 6791–6797 (1982).
22. Petrovich, R. *Geochim. cosmochim. Acta* **45**, 1665–1674 (1981).
23. Gillet, P. & Madon, M. *Bull. Miner.* **105**, 590–597 (1982).

Viral mortality of marine bacteria and cyanobacteria

Lita M. Proctor*† & Jed A. Fuhrman‡

* Marine Sciences Research Center, State University of New York at Stony Brook, Stony Brook, New York 11794-5000, USA

‡ Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-0371, USA

DESPITE the importance of cyanobacteria in global primary productivity¹ and of heterotrophic bacteria in the consumption of organic matter in the sea², the causes of their mortality, particularly the cyanobacteria, are poorly understood. It is usually assumed that mortality is due to protozoan grazing^{3,4} rather than to viral infection, probably because abundances of phage and host in nature are presumed to be low⁵. Previously, either very few marine bacteriophages have been found by plaque assays^{6–9}, or viruses have been simply observed^{10–12} or counted^{13,14} by transmission electron microscopy, with the assumption that 'phage-looking' forms are locally active bacteriophages. Here we report not only high viral abundance in the ocean but also counts of bacteria and cyanobacteria in the final irreversible stage of lytic infection. The latter counts are necessary to evaluate mortality, because the sources, hosts, viability and ages of observed free viruses are unknown; even finding viruses attached to cells does not prove successful infection. Up to 7% of the heterotrophic bacteria and 5% of the cyanobacteria from diverse marine locations contained mature phage; interpretation via culture data indicates that up to 70% of the prokaryotes could be infected. These data demonstrate the existence of a significant new pathway of carbon and nitrogen cycling in marine food webs and have further implications for gene transfer between marine organisms.

† Present address: Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-0371, USA.

We observed a diverse array of virions (Fig. 1), many similar to known bacteriophages, with polygonal heads ranging in diameter from 30 to 160 nm; many had tails which were 50–200-nm long. Virion abundances ranged broadly from 10^6 to 10^{11} per litre of sea water (Table 1). Up to 7% of the heterotrophic bacteria and 5% of the cyanobacteria had mature phage particles inside (Table 2). Mature phage particles within infected bacteria looked similar to free virions. Phage-infected heterotrophic bacteria each contained ~ 10 –100 mature phage particles, whereas phage-infected cyanobacteria often had more (Fig. 2). Our estimates of phage-infected cells and density of phage inside hosts were low because 60–90-nm thin sections contain only $\sim 15\%$

of the contents of cells of average size, and phage particles are often localized within the infected cell (Fig. 2).

The final stage of assembled phage immediately precedes host lysis and represents only a small portion of the time between infection and lysis (latent period)¹⁵. To estimate the portion of the total community infected at any given time, an infection pattern for a marine bacterium, *Cytophaga marinoflava*, whereby

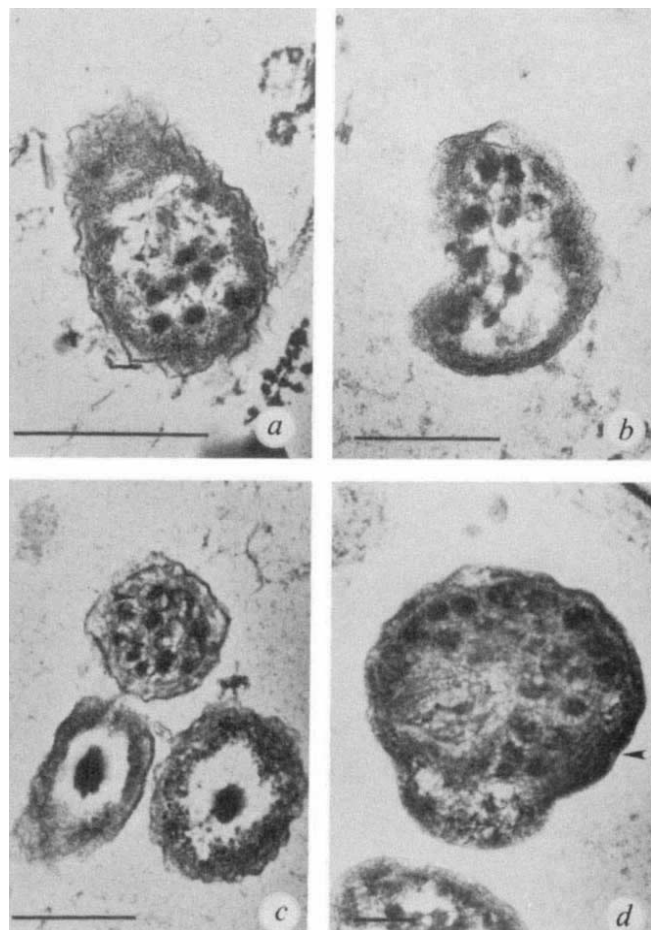


FIG. 1 Electron micrographs of negatively stained phage particles collected by ultrafiltration. *a*, Heterotrophic bacterium with seven associated phage particles from the eastern Caribbean ($12^{\circ}45'N$, $61^{\circ}44'W$; March 1988). Phage heads are 66-nm diameter. Note short tail on phage particles (arrow). *b*, Cluster of tailed phage particles from the eastern Caribbean, March 1988, adjacent to four 88-nm polystyrene latex beads (arrow). Phage heads are 46-nm diameter, tails are 98-nm long. *c*, Tailed phage from Gulf Stream ($22^{\circ}14'N$, $80^{\circ}25'W$; November 1988), adjacent to 88 nm latex bead. Phage head is 92-nm diameter, tail is 68-nm long. *d*, Tailed phage from the eastern Caribbean, March 1988. Note contracted sheath (sh) and exposed tail core (c). Phage head is 89-nm diameter and tail core is 206-nm long. Scale bar, 0.1 μm . Virions were prepared for TEM after concentration at sea by hollow-fibre ultrafiltration²³, except 100,000 molecular mass (MW) cutoff cartridge used, from 40 to 80 litres prefiltered (Gelman A/E, 0.22- μm Durapore) sea water. The sampling scheme took island-effect into consideration; also care was taken not to sample while the ship was emptying bilge and sewage. No surface samples taken. Viral suspension (2 μl) pipetted onto Formvar-coated gold grid, dialysed against 200 μl distilled water for 30 min, dried and negatively stained with 2% uranyl sulphate²⁴. Grids examined on Philips 300 transmission electron microscope at 80 kV.

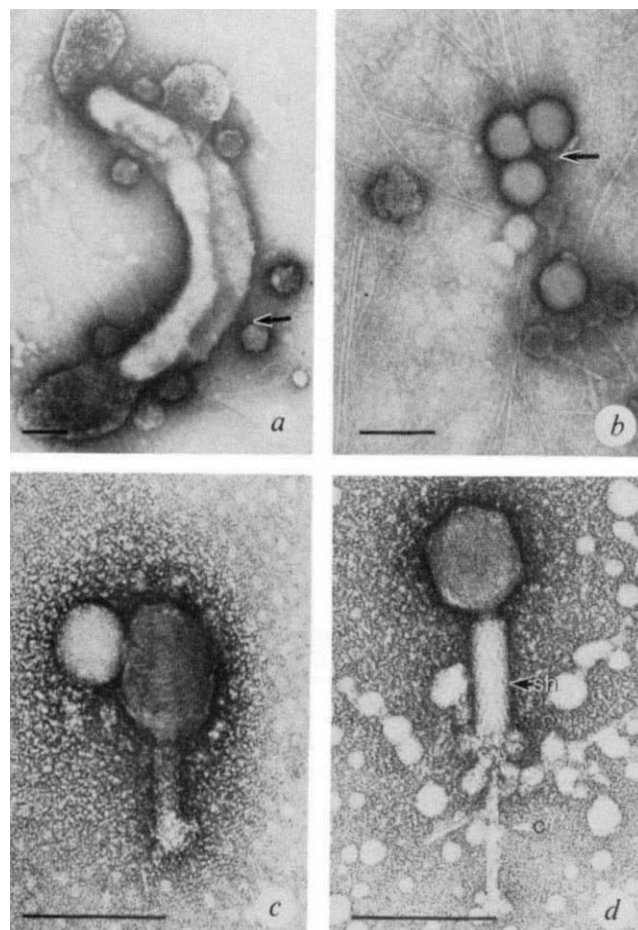


FIG. 2 Electron micrographs of thin sections of cells collected by ultrafiltration. *a*, Heterotrophic bacterium from the Sargasso Sea ($32^{\circ}16'N$, $64^{\circ}37'W$; July 1988), with mature phage particles within the cell. *b*, Heterotrophic bacterium from the western Caribbean ($17^{\circ}38'N$, $80^{\circ}30'W$; November 1988), with mature phage particles forming at one end of the host cell. *c*, Heterotrophic bacteria from the western Caribbean, November 1988, one containing mature phage particles. Note the protein coats of the phage particles are particularly evident (arrow). *d*, Cyanobacterium from the Gulf Stream, November 1988, with mature phage particles beginning to fill the cell. Note thylakoids are partially intact (arrow). Scale bar, 0.5 μm . Bacteria concentrated at sea by ultrafiltration²⁵, except 100,000 molecular mass (MW) cutoff cartridge used, from 100 to 150 litres of prefiltered (Gelman A/E) sea water collected from the same sites as for the virions, then prepared for TEM. Prefiltered sea water was fixed with 0.25% glutaraldehyde to stop metabolic activity of bacteria (inhibition tested with 3H -labelled thymidine incorporation) and prevent enhanced infection due to close contact of host and phage. Recovery efficiency of bacterial cells, determined by enumerating bacteria in whole sea water and bacterial concentrate and corrected for concentration factor and loss of cells to ultrafiltrate, averaged $94 \pm 4\%$. Bacterial concentrate prepared for TEM by covering in agar²⁶, except 2% electron-microscopy grade glutaraldehyde-fixed bacterial concentrate, centrifuged at 45,000g at 4 $^{\circ}C$ for 60 min, covered in 50 μl of 1%-filtered (0.22 μm) 38 $^{\circ}C$ -purified agar, postfixed in 2% osmium tetroxide, dehydrated through graded acetone series, embedded in Epon 812 epoxy resin and sectioned with diamond knife. Gold and silver sections stained with aqueous lead citrate-uranyl acetate and viewed on a Philips 300 transmission electron microscope at 80 kV.

TABLE 2 Distributions of phage-infected heterotrophic bacteria and cyanobacteria

Site	Heterotrophs		Phototrophs	
	% Infected	Cells l ⁻¹	% Infected	Cells l ⁻¹
Long Island Sound	4.1 ± 3.0 (1,000)	6.5 × 10 ⁹	2.8 ± 2.0 (200)	5.0 × 10 ⁸
Eastern Caribbean Sea	3.9 ± 1.3 (2,000)	5.4 × 10 ⁸	1.7 ± 1.1 (400)	2.2 × 10 ⁷
Western Caribbean Sea	2.8 ± 2.1 (1,800)	6.3 × 10 ⁸	1.3 ± 1.1 (300)	4.1 × 10 ⁷
Sargasso Sea	0.9 ± 1.1 (1,200)	4.0 × 10 ⁸	0.8 ± 0.8 (250)	7.0 × 10 ⁶
Gulf Stream	4.3 ± 1.4 (1,400)	6.6 × 10 ⁸	1.0 ± 1.3 (300)	6.0 × 10 ⁷

Distribution of phage-infected heterotrophic bacteria and cyanobacteria in the ocean by TEM; samples collected from same sites as in Table 1. Data expressed as the average percentage and range of phage-infected cells of the total number of cells enumerated at ×20,000 final magnification. Values in parentheses indicate total number of cells enumerated from each site. Cell abundances of heterotrophs, from acridine orange epifluorescence microscopy²¹, and of phototrophs, from autofluorescence microscopy, expressed as cells per litre. Heterotrophic bacteria and cyanobacteria enumerated at ×20,000 final magnification with presence of mature phage particles within cells scored as infected. Cyanobacteria can also contain carboxysomes, polyphosphate bodies and other inclusions that can be mistaken for phage. Staining characteristics of phage, arrangement of phage within cell and total number of phage were all used to distinguish between metabolic byproducts and phage particles²².

the final visible stage of assembled phage represents 10% of the latent period¹⁶, was applied to our field data. This indicated that although an average of 3.2% of heterotrophic bacteria and 1.5% of cyanobacteria contain mature phage, about 32% and 15% respectively, are infected at any given time.

To estimate absolute mortality rates for comparison with growth, the relationship between latent periods and host division rates must be known. Although such data are rare, for a marine vibrio it seems that the latent period is about equal to the doubling time of the uninfected host¹⁷. If this holds for other naturally occurring organisms, and cell division is balanced by total mortality (that is, the steady-state where half the daughter cells survive to divide again), then the per cent of total mortality attributable to viruses is twice the total per cent of cells infected in the microbial community at any given time. For example, if 50% of the bacteria are lysed by viruses and the other 50% survive to divide, then viruses account for 100% of total bacterial mortality; alternatively, if 5% of the bacteria are lysed by viruses, 45% of the bacteria are killed by other agents and 50% survive to divide, then 10% of total bacterial mortality is due to viral lysis. Although this model yields only a first approximation, it demonstrates that viral infection can be a significant source of prokaryote mortality, responsible for a large but variable fraction of the total—of the order of 30% for the cyanobacteria and 60% for the heterotrophic bacteria. Preliminary independent estimates of bacterial cell loss in sea water with added concentrated viruses are consistent with this model¹⁸. In addition, a previous study showed significant (~50% of total) bacterial mortality in 0.6-µm filtrates essentially free of visible protozoa¹⁹; although

in this study antibiotic sensitivity implicated eukaryotes as the cause of mortality, the authors indicated that viruses could have been the cause. Improved estimates of total mortality will require additional information about the relationship between the rates of host division and viral lysis.

We have shown that bacteriophage infection could be a significant mechanism of mortality in marine prokaryotes. These mortality mechanisms are critical in understanding the fate of both primary and secondary marine production. Cell lysis releases cellular constituents and could constitute a significant new pathway of C and N cycling for both primary producers and consumers. Thus cell lysis is relevant to our understanding of geochemical cycles in the sea and therefore to related topics such as global warming. Even unsuccessful viral infection could be important in C and N cycling, because inactivated phage are a source of nucleic acids and amino acids for bacteria. Viral infection also has implications for genetic exchange in the sea. The abundant virions in the world's oceans have the potential for significant transduction, and should be included in any study of the release of genetically engineered bacteria. Although we do not yet understand the temporal or spatial scales at which host-virus encounter occurs, viral mortality must be considered in models of marine systems. □

TABLE 1 Distribution of virions

Site	Date	Depth (m)	Free viral particles (10 ⁹ l ⁻¹)	N
Long Island Sound	July 1987	1	150 ± 95.0	4
Eastern Caribbean Sea	March 1988	5–200	1.9 ± 1.28	9
Western Caribbean Sea	November 1988	5–200	4.8 ± 3.01	7
Sargasso Sea	July 1988	25	0.003 ± 0.0015	2
Gulf Stream	November 1988	50	460 ± 28.6	2

Distribution of virions in the ocean by transmission electron microscopy (TEM), collected from several stations at each site. Cruise tracks for eastern Caribbean Sea: 20°08'N, 68°23'W to 12°35'N, 61°45'W; for western Caribbean Sea: 18°54'N, 76°34'W to 10°36'N, 82°26'W; stations for Long Island Sound: 40°55'N, 73°09'W; for Sargasso Sea: 32°15'N, 64°35'W; for Gulf Stream: 22°45'N, 82°16'W. Average abundance and range of virions per litre of sea water after correction for concentration and recovery efficiency; N, number of stations. Virions counted, after negative staining, with 88-nm polystyrene latex beads added as internal marker at known concentration to viral concentrate²⁰. Entire grid examined at ×15,000 final magnification with beads, recognizable bacteriophage heads and tails counted. Final concentration of virions related to latex bead abundance, concentration factor and recovery efficiency. Recovery efficiency, determined by adding known quantity of T5 bacteriophage to seawater samples and enumerating T5 phage with indigenous virions and latex beads, was 81% ± 12% (n = 3).

Received 17 July; accepted 15 November 1989.

1. Waterbury, J. B., Watson, S. W., Valois, F. W. & Franks, D. G. in *Phytoplankton and Picoplankton* (eds Platt, T. & Li, W. K. W.) 71–120 (Department of Fisheries and Oceans, Ottawa, Canada, 1986).
2. Azam, F. *et al. Mar. Ecol. Prog. Ser.* **10**, 257–263 (1983).
3. Fenchel, T. *Mar. Ecol. Prog. Ser. Microbiol.* **9**, 35–42 (1982).
4. Sherr, E. B. & Sherr, B. F. *Nature* **325**, 710–711 (1987).
5. Wiggins, B. & Alexander, M. *Appl. env. Microbiol.* **49**, 19–23 (1985).
6. Spencer, R. in *Symposium on Marine Microbiology* (ed. Oppenheimer, C. H.) 350–365 (Charles Thomas, Springfield, Illinois, 1963).
7. Kriss, A. E. in *Microbial Populations of Oceans and Seas* (eds Kriss, A. E., Mishustina, I. E., Mitskevich, N. & Zentsova, E.) 230–237 (St Martin's, New York, 1967).
8. Hidaka, T., Kawaguchi, T. & Shirahama, M. *Mem. Fac. Fish. Kagoshima Univ.* **28**, 47–55 (1979).
9. Moebus, K. *Helgoländerwiss. Meeresunters.* **36**, 375–391 (1983).
10. Sieburth, J. McN. *Sea Microbes*, 62 (Oxford University Press, New York, 1979).
11. Johnson, P. W. & Sieburth, J. McN. *J. Phycol.* **18**, 318–327 (1982).
12. Frank, H. & Moebus, K. *Helgoländerwiss. Meeresunters.* **41**, 385–414 (1987).
13. Torrella, F. & Morita, R. Y. *Appl. env. Microbiol.* **37**, 774–778 (1979).
14. Bergh, Ø., Barsheim, K. Y., Bratbak, G. & Haldal, M. *Nature* **340**, 467–468 (1989).
15. Kellenberger, E., Sechaud, J. & Ryter, A. *Virology* **8**, 478–490 (1959).
16. Valentine, A. F. & Chapman, G. B. *J. Bact.* **92**, 1535–1554 (1966).
17. Smith, L. S. & Krueger, A. P. *J. gen. Physiol.* **38**, 161–168 (1954).
18. Proctor, L. M., Fuhrman, J. A. & Ledbetter, M. C. *Eos* **69**, 1111–1112 (1988).
19. Fuhrman, J. A. & McManus, G. B. *Science* **224**, 1257–1260 (1984).
20. Miller, M. in *Principles and Techniques of Electron Microscopy* (ed. Hayat, M. A.) 89–128 (Van Nostrand, New York, 1974).
21. Hobbie, J. E., Daley, R. & Jasper, J. P. *Appl. env. Microbiol.* **33**, 1225–1228 (1977).
22. Leach, J. A., Lee, K. W., Benson, R. L. & Martin, E. L. *J. Physiol.* **16**, 307–310 (1980).
23. Bergh, Ø., Rotem, Y. & Katznelson, E. *Water Res.* **10**, 279–284 (1976).
24. Estis, L. F., Haschemeyer, R. H. & Wall, J. S. *J. Microscopy* **124**, 313–316 (1981).
25. Sherr, B. F., Sherr, E. B. & Fallon, R. D. *Appl. env. Microbiol.* **53**, 958–965 (1987).
26. Johnson, P. W. & Sieburth, J. McN. *Limnol. Oceanogr.* **24**, 928–935 (1979).

ACKNOWLEDGEMENTS. We thank Dr M. C. Ledbetter and Mr N. Tempel for discussion and assistance with TEM, Drs E. J. Carpenter and D. Capone for shipment on the R/V *Columbus Iselin*, Drs M. C. Ledbetter, J. G. Mitchell and F. Torrella for comments, and Dr J. McN. Sieburth for discussion about marine bacteria. This work was supported by the NSF and DOE.

Morphogenetic basis for phenotypic differences in hydroid competitive behaviour

Leo W. Buss* and Richard K. Grosberg†

* Departments of Biology and Geology and Geophysics, Yale University, New Haven, Connecticut 06511, USA

† Department of Zoology, University of California, Davis, California 95616, USA

FOR many plants, fungi and sessile invertebrates, the availability of habitable space often limits their size, reproductive output and survival. Intraspecific and interspecific variation in the ability to compete for space is common^{1–3}, but the way in which competitive ability is regulated remains largely unknown. We describe here a system of intraspecific competition in the hydroid *Hydractinia symbiolongicarpus* in which the principal mechanisms underlying variation in competitive ability have proven amenable to experimental analysis. Competitive interactions can be aggressive, involving the induction of a specialized fighting organ called a hyperplastic stolon, or non-aggressive, in which no hyperplastic stolons are produced. Colonies display continuous variation in the ontogenetic appearance of tissues that differ in their competence to be induced to produce hyperplastic stolons. We find that the outcome and kind of competitive interaction between strains are predictable given knowledge of their morphologies. In this hydroid, complex competitive behaviour arises from a coupling of discrete morphogenetic potential of differing tissue types with continuous variation in the ontogenetic and astogenetic appearances of these tissues.

Colonies of *H. symbiolongicarpus* are commonly found encrusting gastropod shells inhabited by pagurid hermit crabs⁴. Intraspecific competition occurs at high frequency in natural populations^{5–7}. Encounters between allogeneic colonies (that is, interactions between unrelated genotypes) result in one of two different rejection responses^{8–11}. The first response—passive rejection—is characterized by the secretion of a fibrous matrix by both colonies¹⁰, and is accompanied by a virtual cessation of growth along the border between the two colonies (Fig. 1a). Aggressive rejection (Fig. 1b), by contrast, is mediated by the induction of a specialized organ of aggression, the hyperplastic stolon^{10–14}. On encountering foreign tissues, stolons hypertrophy as a result of widespread movement of nematocytes into the stolon tip and subsequent nematocyst discharge^{10,14} (Fig. 1d). Nematocyte recruitment and discharge continues until either the foreign tissue is destroyed or until the hyperplastic stolon is destroyed by a similar hyperplastic response from its neighbour.

Colonies of *H. symbiolongicarpus* display considerable variation in gross colony morphology during early ontogeny^{15–16}. Colonies grow by the asexual iteration of feeding polyps, by the radial expansion of a basal ectodermal mat, and by the peripheral extension of tubular stolons that ramify and anastomose over the substratum. Colonies range from completely stolonless (Fig. 2a) to highly stoloniferous (Fig. 2b). Experiments in which clonal replicates of each of 14 different *H. symbiolongicarpus* strains were reared under constant conditions in the laboratory (that is, 'common garden experiments'), revealed significant variation among the strains in the relative rates of mat, polyp, and stolon production (Table 1).

In addition to ontogenetic variation in relative rates of mat, polyp, and stolon production, there is considerable variation in the form of mat tissue at the colony periphery. The ectoderm of the mat tissues can consist of unfused tubes (Fig. 2c) similar to that found in stolons; alternatively, the ectoderm of adjacent gastrovascular canals can be laterally fused into a continuous sheet. In some colonies, all mat tissue consists of fused ectoderm,

whereas in other colonies, there is a central region of fused mat surrounded by a peripheral region of unfused mat. This region of unfused mat is called the astogenetic zone¹⁷. Thus, a zone of astogeny is entirely absent in some colonies, whereas an astogenetic zone of varying width occurs in other colonies. As for ontogenetic variation, common garden experiments suggest that there is a substantial genetic component to variation for both the presence or absence of an astogenetic zone, as well as for the width of the astogenetic zone (Table 1).

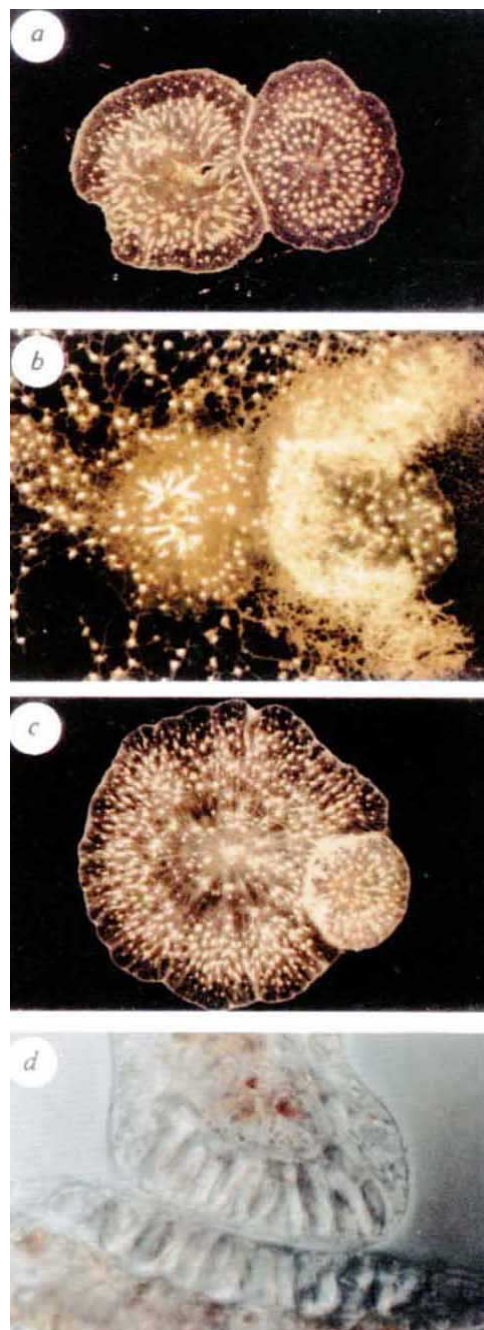


FIG. 1 a, A passive rejection between two stolonless *Hydractinia* colonies. b, An aggressive rejection between two stoloniferous *Hydractinia* colonies. Note the extensive development of hyperplastic stolons separating the two colonies. c, An interaction between two stolonless colonies escalating from an initially passive response to an aggressive response. Note the development of hyperplastic stolons along the margin of the larger colony in contact with the smaller colony. d, A Nomarski interference micrograph of contact between allogeneic stolons showing the array of nematocysts at the point of contact (photograph by R. Lange and G. Plickert).

As a consequence of ontogenetic and astogenetic variation in colony morphology, competitive encounters between colonies of *H. symbiolongicarpus* can involve any combination of three different tissues: stolons, unfused ectodermal mat (termed UM), or fused ectodermal mat (termed FM). Allogeneic encounters between stolons and any other tissue type result in the induction of a hyperplastic response⁸. A series of grafting experiments clarifies the morphogenetic potential of the two types of mat tissues. Reciprocal transplants of central FM tissues from one colony into contact with either the central FM ($n=15$) and peripheral UM ($n=15$) of an allogeneic colony yielded a passive response in all cases. By contrast, grafts from the peripheral UM of the same colonies in contact with either central FM ($n=15$) or peripheral UM ($n=15$) initially showed a passive response, followed within 1–3 weeks by the induction of an aggressive response. This aggressive response began as stolons emerged from the acellular matrix at the periphery of the mat; these subsequently differentiated into hyperplastic stolons. Thus, the two forms of rejection—passive and aggressive—are determined by differing morphogenetic potentials of the tissue types coming into contact with one another.

The discrete developmental potentials of stolon and mat tissues, when coupled with strain-specific variation in the pattern and rate of appearance of these tissues, potentially leads to a complex array of competitive behaviours for this simple diploblastic organism. Stoloniferous colonies should behave aggressively (Fig. 1b) and stolonless colonies with fused mat should not (Fig. 1a). But a range of more complex behaviours are also possible. For example, stolonless colonies with unfused mat

should initially behave non-aggressively, only to escalate interactions to involve aggression (Fig. 1c). Conversely, an aggressive response can be de-escalated if, during the course of a competitive encounter, the central fused mats of one, or both, colonies come into contact.

Competitive ability and outcome should thus be determined by the ontogenetic and astogenetic state of the competing colonies. We examined the outcome of pairwise size-symmetric competitive interactions between colonies in laboratory culture. We established these interactions by explanting equally sized explants of two colonies, separated by ~ 0.5 cm, onto glass microscope slides. These interactions were monitored roughly fortnightly until one colony had eliminated the other, or in cases of prolonged interactions, for nearly six months.

Stoloniferous versus stoloniferous phenotypes: We established replicated competitive interactions between seven stoloniferous

TABLE 1 Common garden experiments

Series	Tissue type	Degrees of freedom	F	Significance level (P)
Stoloniferous	Polyps	18, 78	23.43	0.001
	Mat	18, 81	17.59	0.001
	Stolon	18, 81	21.93	0.001
Stolonless	Polyps	24, 136	8.15	0.001
	Mat	24, 136	14.30	0.001
	Astogeny	12, 66	12.27	0.001

Analysis of variance for the effects of colony genotype on growth trajectories of polyp, mat, stolon and width of the astogenetic zone in two common garden experiments. These experiments were established by clonally propagating field collected colonies. A portion of ectodermal mat containing 3–5 feeding polyps (that is, an explant) was surgically removed, placed on a glass microscope slide, and held in place with a loop of thread. Explants attached to slides within 1–7 days and threads were removed. Stock colonies established in this manner were maintained in recirculating aquaria and fed to repletion daily with brine shrimp nauplii. Common garden experiments were established from stock colonies in two series. In the first series, five replicates of each of seven stoloniferous strains were explanted onto glass slides, maintained as were stock colonies, and their growth monitored at roughly 2-week intervals by macrophotography. In the second series, the growth of 4–6 replicates of each of seven stolonless strains was monitored. For each sampling interval, the number of polyps and the area of mat was calculated from photographs. In addition, stolon area (that is, the area prescribed by connecting the endpoints of adjacent stolon tips) was calculated for all stoloniferous strains and the width of the astogenetic zone measured for all stolonless strains. Mat and stolon area were calculated by tracing the outline onto a digitizer interfaced to a microcomputer. Log-transformed data were analysed as a two-factor design with repeated measures on one factor¹⁹. The main effects are strain and day, where day is treated as a categorical variable and each replicate was observed under all levels of day. When testing between-strain differences in growth rate, the relevant null hypothesis is that there is no strain-by-day interaction. This interaction mean square is tested against the day-by-replicate within-strains error mean square. Common garden experiments of this sort provide an estimate of broad-sense heritability²⁰. Quantitative genetic analysis of variation in colony morphology reveals low additive to non-additive genetic variation for shape characters²¹.

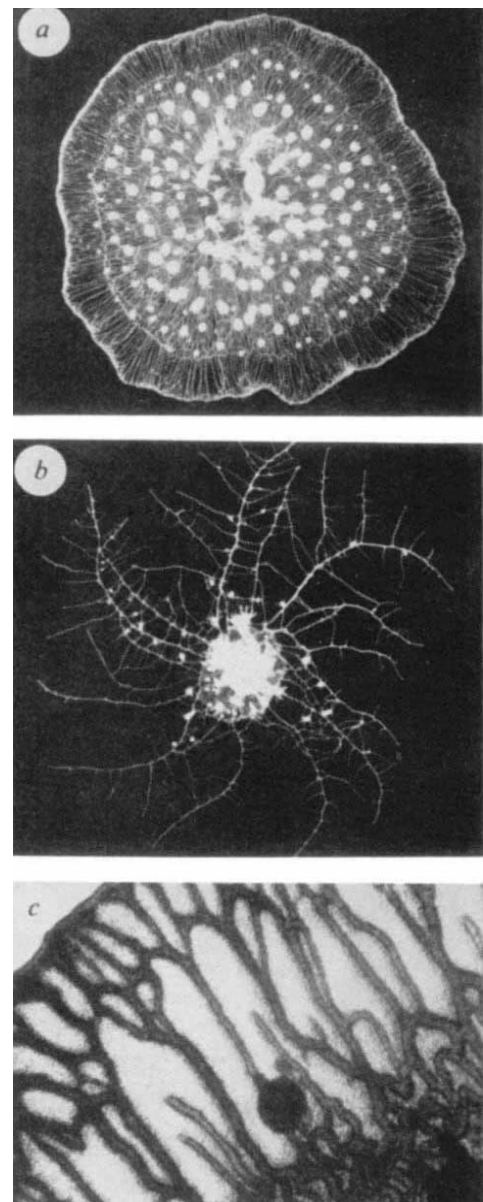
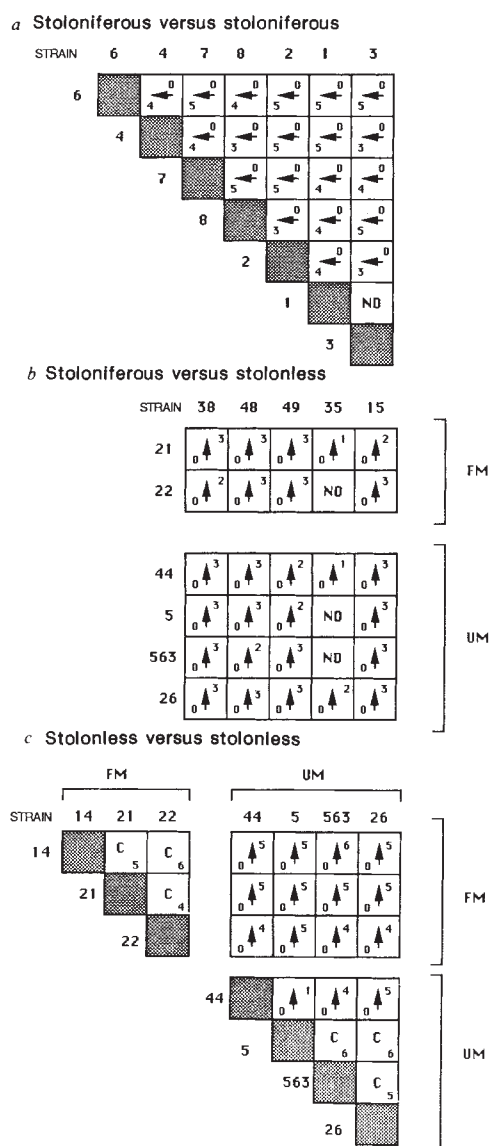


FIG. 2 a, A stolonless colony. b, A stoloniferous colony. c, The astogenetic zone of a UM colony, showing peripheral gastrovascular canals embedded within a transparent matrix. Note that beyond this marginal zone, gastrovascular canals become more dense. In central regions of the colony, the ectoderm of adjacent canals fuse to form a continuous sheet (not shown).

strains in 20 of the 21 possible pairwise combinations and observed these interactions for up to 173 days. All pairwise interactions ($n=86$) were characterized by the reciprocal development of hyperplastic stolons. In 95% of all interactions one colony killed the other within 37–150 days. Interactions between stoloniferous colonies were transitive, with rapidly growing strains overcoming the slowly growing strains in all pairwise combinations (Fig. 3a).



Stolonless versus stolonless phenotypes: Replicated interactions between seven stolonless strains were established for three UM and four FM strains in all pairwise combinations and observed for up to 113 days. No evidence of aggression was observed in any of the three pairwise combinations involving FM strains ($n=15$). All FM colonies in competition with other FM colonies coexisted for the entire experimental period (Fig. 3b). In contrast to the results of FM versus FM interactions, all pairwise combinations between FM and UM strains ($n=48$) resulted in either the elimination of the FM colony (44%, $n=21$) or in the loss of considerable tissue mass by the FM colony at the termination of the experiment (Fig. 3b). In each case, the UM colony began the interaction with a passive response, followed after 1–3 weeks by an aggressive response as stolons emerged from the acellular matrix at the edge of the mat. These emergent stolons soon became hyperplastic. By comparison, no FM colonies behaved aggressively.

Interactions between UM strains and other UM strains resulted in coexistence in three combinations ($n=17$) and prolonged aggression in three other combinations ($n=10$) (Fig. 3b). In all cases of coexistence, an initially passive response was followed in 1–3 weeks by aggression. After an aggressive interval of varying duration, a passive response spread along the interactive margin as the central FM of both colonies came into tissue contact. All cases of prolonged aggression occurred in interactions involving a single slowly growing strain (Fig. 3b). This strain was either eliminated, or in the process of being eliminated, before the central FM regions came into contact.

Stolonless versus stoloniferous phenotypes: Interactions between five stoloniferous strains and six stolonless strains (2 FM, 4 UM) were established in 27 of the 30 possible pairwise combinations and observed for 45 days. In all interactions involving a stolonless FM colony ($n=23$), the stoloniferous colony was eliminating the FM colony in all replicates at the termination of the experiment. All stoloniferous colonies produced hyperplastic stolons and all FM colonies did not. In all interactions involving a UM colony ($n=48$), the stoloniferous colony produced hyperplastic stolons on contact with the UM colony, whereas the UM colony initially displayed a passive response. Eventually UM colonies displayed a hyperplastic response; nevertheless, the stoloniferous colony was overgrowing the UM colony in all interactions at the termination of the experiment.

These experiments demonstrate that competitive ability is a simple deterministic function of growth morphology in this hydroid. Despite their simplicity, hydractiniid hydroids display a repertoire of aggressive behaviours like those of behaviourally sophisticated organisms¹⁸. Some colonies always behave aggressively, others never behave aggressively, and yet others can switch their behaviour from aggression to non-aggression (or vice versa). The capacity to display a diverse array of competitive behaviours is attributable to the central part played by the developmental processes of induction and competence in this system. Here competitive ability is not coded as a series of discrete alternatives; indeed morphological variation seems to be continuous. This variation, nevertheless, results in discrete competitive phenotypes because tissue types differ in their competence to be induced to form hyperplastic stolons. Differing competence defines a series of discrete behaviours for differing tissues. As such, an enormously complex range of competitive phenomena can occur as a consequence of discrete morphogenetic potential of differing tissue types whose ontogenetic appearance displays continuous variation. □

Received 17 May; accepted 24 October 1989.

1. Jackson, J. B. C. in *Biotic Interactions in Recent and Fossil Benthic Communities* (eds Tevesz, M. J. S. & McCall, P. L.) 39–120 (Plenum, New York, 1983).
2. Buss, L. W. in *Community Structure* (eds Case, T. & Diamond, J.) 517–536 (Harper and Row, New York, 1986).
3. Keddy, P. *Competition* (Chapman and Hall, London, 1989).

FIG. 3 Results of pairwise competitive interactions between hydroid colonies differing in ontogeny and astogeny. **a**, Interactions between seven stoloniferous strains. Each cell in this matrix represents interactions between the strain on the left and the strain above. The arrow points toward the strain which eliminated or was judged to be in the process of eliminating the other at the termination of the experiment. The number in the bottom left-hand corner of each cell indicates the number of replicates won by the colony on the left, and the number in the upper right-hand corner indicates the number of replicates won by the colony above. Strains are ordered left to right by their relative growth rates as measured in common garden experiments (that is, strains were ranked for log-transformed growth rate for each tissue type and the sum of the ranks for all three tissue types used to order colonies). ND, no data. **b**, Interactions between six stolonless strains (rows) and five stoloniferous strains (columns). C, coexistence. Strains are categorized by their astogenetic state. Number of replicates indicated in bottom left-hand corner of each cell. All other notations as in **a**. **c**, Interactions between seven stolonless strains. All other notations as in **b**.

4. Buss, L. W. & Yund, P. O. *J. mar. biol. Ass. UK* **69** (in the press).
5. Yund, P. O., Cunningham, C. W. & Buss, L. W. *Ecology* **68**, 971–982 (1987).
6. Buss, L. W. & Yund, P. O. *Ecology* **69**, 646–654 (1988).
7. Yund, P. O. & Parker, H. M. *J. exp. mar. Biol. Ecol.* **125**, 63–82 (1989).
8. Hauenschild, C. *Wilhelm Roux. Arch. EntwMech. Org.* **147**, 1–41 (1954).
9. Hauenschild, C. *Z. Naturforsch.* **11**, 132–138 (1956).
10. Buss, L. W., McFadden, C. S. & Keene, D. R. *Biol. Bull.* **167**, 139–158 (1985).
11. Buss, L. W. & Shenk, M. A. in *Defense Molecules* (eds Marchionis, J. J. & Reinisch, C.) (A. R. Liss, New York) (in the press).
12. Müller, W. A. *Wilhelm Roux. Arch. EntwMech. Org.* **155**, 181–268 (1964).
13. Ivker, F. B. *Biol. Bull.* **143**, 162–174 (1972).
14. Lange, R., Plickert, G. & Müller, W. A. *J. exp. Zool.* **249**, 284–292 (1989).
15. McFadden, C. S., McFarland, M. J. & Buss, L. W. *Biol. Bull.* **166**, 54–67 (1984).
16. Blackstone, N. W. & Yund, P. O. *Paleobiology* **15**, 1–10 (1989).
17. Boardman, R. S. & Cheetam, A. H. in *Animal Colonies* (eds Boardman, R. S., Cheetam, A. H. & Oliver, W. A. Jr) 121–222 (Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania, 1973).
18. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, 1982).
19. Winer, B. J. *Statistical Principles in Experimental Design* (McGraw-Hill, New York, 1971).
20. Falconer, D. S. *Introduction to Quantitative Genetics* (Longman, New York, 1981).
21. Blackstone, N. W. & Buss, L. W. *Evolution* (in the press).

ACKNOWLEDGEMENTS. We thank M. Bertness, N. Blackstone, J. Bonner, J. Buss, D. Carlson, R. Crozier, C. Cunningham, A. Ellison, D. Green, G. Hartnett, N. Knowlton, C. McFadden, A. Shenk, H. Waldeman, J. R. Vaisnys and P. Yund for discussions of this work, and T. Jenkins, R. Lerner, M. Marthas, J. Taschner and H. Waldman for technical assistance. This work was supported by the NSF.

Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*

Michael J. Ryan*, James H. Fox†, Walter Wilczynski‡ & A. Stanley Rand‡

* Department of Zoology and † Department of Psychology, University of Texas, Austin, Texas 78712, USA

‡ Smithsonian Tropical Research Institute, Balboa, Panama

THE sensory bases of species and population mate preferences are well known^{1–3}; in frogs properties of the female auditory system influence such preferences^{2,3}. By contrast, there is little understanding of how sensory characteristics could result in sexual selection within a population. One possible mechanism is that females are more sensitive to male courtship signals that deviate from the population mean. We document this mechanism in the frog *Physalaemus pustulosus*. Female basilar papilla tuning is biased toward lower-than-average frequencies in the 'chuck' portion of the male's call, explaining female preference for the lower-frequency chucks produced by larger males. The tuning does not differ between *P. pustulosus* and its close relative *P. coloradum*, a species in which males never evolved the ability to produce chucks; thus the female tuning evolved before the chuck and therefore the chuck played no role in the evolution of the preference. This allows us to reject two popular hypotheses for the evolution of this female preference (runaway sexual selection and natural selection) in favour of a third: sexual selection for sensory exploitation.

Male *P. pustulosus* produce an advertisement call consisting of a 'whine', which is necessary and sufficient for species recognition, followed by 0–6 'chucks', which increase the attractiveness of the call to females^{4–8}. Chucks have a fundamental frequency of ~220 Hz with 15 harmonics. Most of the energy ($\bar{X}$ = 90%, s.e. = 0.02, N = 110) is above 1.5 kHz, and thus mainly stimulates basilar papilla receptors rather than the lower-frequency amphibian papilla receptors^{9,10}. In phonotaxis experiments, females prefer synthetic calls having chucks with lower fundamental frequencies. Larger males produce lower-frequency chucks and have greater mating success in nature. Thus preferential female phonotaxis generates sexual selection^{4–8}.

This preference could result from the upper harmonics of lower-frequency calls eliciting greater neural stimulation of the high-frequency basilar papilla receptors due to differences in peak frequency or harmonic structure¹¹. Similar mechanisms appear to guide interspecific and interpopulational mate

choice^{3,10,12}. In all studies to date in which female mate preferences based on call frequency are expressed, tuning of the peripheral auditory system predicts the preference by indicating frequencies that would most strongly stimulate the auditory system. There are no reported examples of any mechanisms overriding this peripheral bias. We tested whether differences in auditory stimulation could account for interspecific mate choice by obtaining acoustically-evoked multiunit activity from the torus semicircularis (inferior colliculus) to estimate the frequency sensitivities of the amphibian and basilar papillae^{3,13}. Resulting audiograms were averaged and truncated below 1.5 kHz to yield a mean basilar papilla audiogram (Fig. 1). We then determined the Fourier spectrum (energy versus frequency)³ of chucks from 54 males randomly sampled from a population on Barro Colorado Island, Panama^{4–8} (Fig. 1). The mean dominant (peak energy) frequency of the chucks (2.55 kHz) was higher than the mean best frequency of the basilar papilla (2.13 kHz).

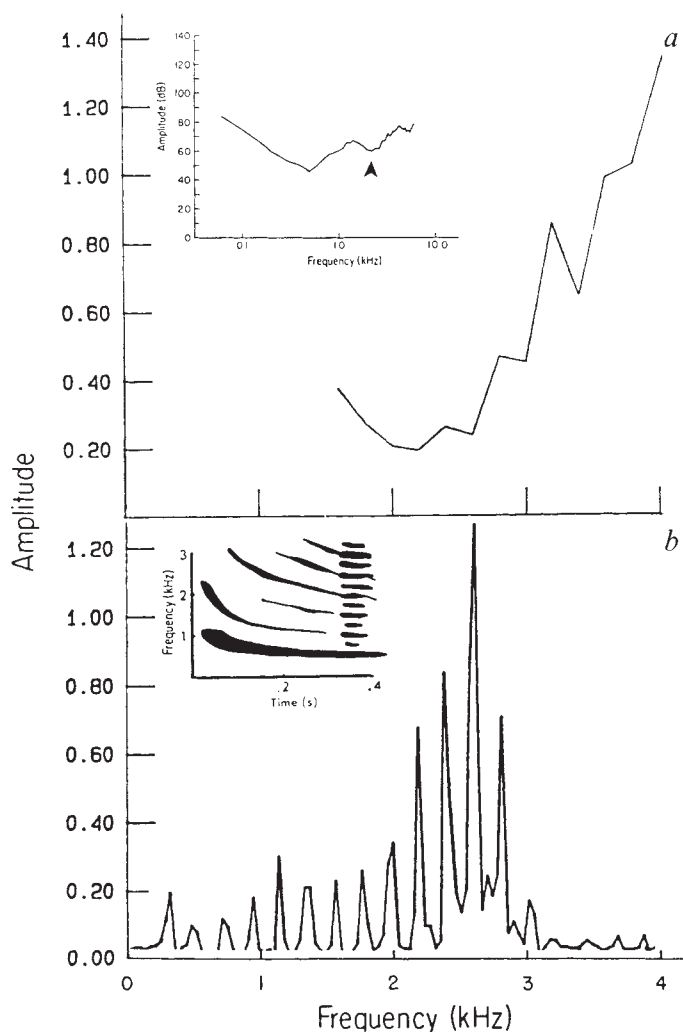


FIG. 1 *a*, The mean audiogram of the basilar papilla of *P. pustulosus* derived from five individuals. Audiograms represent thresholds as a function of frequency, determined for sinusoidal, closed-field stimuli using 1–2 M Ω glass electrodes. The truncation of the audiogram below 1.5 kHz to eliminate influences of amphibian papilla neurons and the slight broadening of the tuning curve resulting from averaging biases the results toward the null hypothesis. Insert, audiogram from a single frog; basilar papilla best frequency is marked by arrow. Male and female audiograms did not differ. *b*, Representative Fourier spectrum of a chuck. Insert, sonogram of a whine plus a chuck.

We quantified each chuck's stimulation of an average basilar papilla using a computer model that treated the audiogram as a filter and measured how much energy in a chuck would pass through it. We then determined how a change in the frequencies of each chuck would influence the amount of neural stimulation. Each chuck's Fourier spectrum was multiplied by a series of numbers ('frequency multipliers') in 0.01 increments between 0.95 and 1.05, and resultant spectra corrected so that total power remained constant. For the chuck of each male, we determined the optimal frequency multiplier, that is, one resulting in a chuck that elicited the greatest neural stimulation (the 'optimal chuck').

Our hypothesis that chucks with lower fundamental frequencies elicit greater neural stimulation predicts that the average optimal frequency multiplier is less than 1.0 (the unmanipulated chuck). Our data support this hypothesis: the average optimal frequency multiplier is 0.98 (s.e. = 0.009), and is significantly less than 1.0 (one-tailed Student's *t* test; $t = 2.0$, d.f. = 53, $P = 0.025$; Fig. 2). Also consistent with our hypothesis, the mean dominant frequency of the optimal chucks ($\bar{X} = 2.489$ kHz, s.e. = 0.05) is significantly lower than that of the unmanipulated chucks ($\bar{X} = 2.552$ kHz, s.e. = 0.058; one-tailed paired *t* test; $t = 2.48$, d.f. = 53, $P = 0.02$). Therefore, the patterns of mate choice reported in nature⁴⁻⁷ and the response of females in phonotaxis experiments⁴⁻⁷ are predicted by the relationship

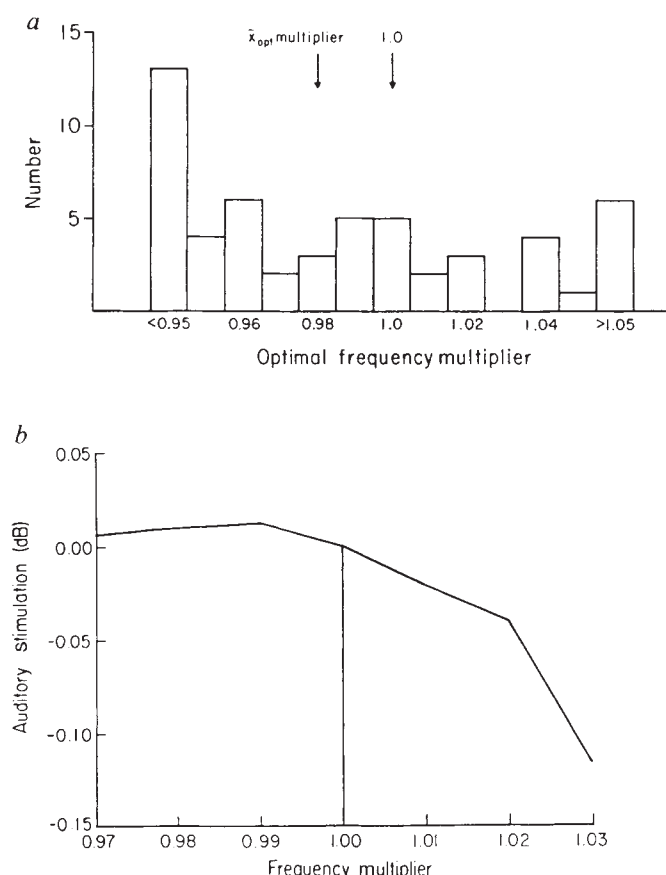


FIG. 2 a, The distribution of frequency multipliers resulting in chucks that maximize stimulation of the basilar papilla. Stimulation by a call was assessed by summing call pressure squared divided by auditory threshold squared at 10-Hz intervals through the basilar papilla audiogram, then taking $20 \times \log$ of the square root of this sum to yield the energy passing through the basilar papilla filter function. b, Average amount of neural stimulation, in dB, is $20 \times \log$ of the ratio of energies passing through an average basilar papilla audiogram from a manipulated chuck and an unmanipulated chuck, as a function of frequency multiplier across all calls tested.

between peripheral tuning and chuck frequencies. Furthermore, the disadvantage of producing a nonoptimal chuck is asymmetric (Fig. 2). As the chuck falls below optimal frequency, neural stimulation decreases only slightly, but there is a greater decrement in stimulation when the frequency of the chuck is increased above the optimal chuck. The asymmetry indicates stronger selection against chucks that are too high relative to chucks that are too low in frequency.

Three alternative hypotheses address the evolution of female preferences. Fisher's runaway sexual selection hypothesis¹⁴⁻¹⁶ suggests that female preferences and male traits are genetically correlated, and that preferences will evolve as a correlated response to selection on the male trait. The natural selection hypothesis^{17,18} postulates that female preferences evolve due to reproductive benefits accrued by choosing certain males. Sexual selection for sensory exploitation^{11,19-22} hypothesizes that males evolve traits that exploit preexisting biases in the female's sensory system. Only the sensory exploitation hypothesis predicts that the evolution of the female trait (basilar papilla tuning) precedes the male trait (the chuck). Fisherian and natural selection hypotheses require the male trait to precede or coincide with the female trait. We assessed these hypotheses by measuring basilar papilla tuning in seven individuals of the closely related species *P. coloradum*. This species does not produce chucks, and the ability to produce chucks was derived in *P. pustulosus* after these species diverged^{23,24}.

There was no significant difference between the average basilar papilla best frequencies of *P. pustulosus* ($\bar{X} = 2.13$ kHz, s.e. = 0.07) and *P. coloradum* ($\bar{X} = 2.23$ kHz, s.e. = 0.13; two-tailed Student's $t = 0.63$, $P = 0.54$; Bartlett test for homogeneity of variances, $F = 2.31$, d.f. = 1, $P = 0.13$). It is most parsimonious to assume that the identical basilar papilla tuning in the two species is a trait shared by a common ancestor, and therefore existed before the chuck evolved in *P. pustulosus*. Female *P. pustulosus* do have more eggs fertilized when they mate with larger males⁴. Although this advantage could help maintain the preference once it is expressed, it cannot be argued that females evolved their tuning characteristics to gain this reproductive advantage. Rather, the existence of identical tuning before and after the evolution of the chuck supports the sensory exploitation hypothesis, that in *P. pustulosus* lower frequency chucks exploit preexisting biases in the female's sensory system, but not the Fisherian or natural selection hypotheses. □

Received 30 May; accepted 26 October 1989.

1. Alcock, J. *Animal Behaviour* (Sinauer, Sunderland, 1988).
2. Ryan, M. J. *Am. Sci.* (in the press).
3. Ryan, M. J. & Wilczynski, W. *Science* **240**, 1786-1788 (1988).
4. Ryan, M. J. *Science* **209**, 523-525 (1980).
5. Ryan, M. J. *Evolution* **37**, 261-272 (1983).
6. Ryan, M. J. *J. comp. Physiol.* **150**, 217-221 (1983).
7. Ryan, M. J. *The Tungara Frog, A Study in Sexual Selection and Communication* (University of Chicago Press, 1985).
8. Rand, A. S. & Ryan, M. J. *Z. Tierpsychol.* **57**, 209-214 (1981).
9. Wilczynski, W. & Capranica, R. R. *Prog. Neurobiol.* **22**, 1-30 (1984).
10. Zakon, H. H. & Wilczynski, W. in *The Evolution of the Amphibian Auditory System* (eds Fritzsche, B., Ryan, M. J., Wilczynski, W., Hetherington, T. E. & Walkowiak, W.) 125-155 (Wiley, New York, 1988).
11. Ryan, M. J. & Rand, A. S. *Evolution* (in the press).
12. Fritzsche, B., Ryan, M. J., Wilczynski, W., Hetherington, T. E. & Walkowiak, W. eds *The Evolution of the Amphibian Auditory System* (Wiley, New York, 1988).
13. Walkowiak, W., Capranica, R. R. & Schneider, H. *Behav. Processes* **6**, 223-237 (1981).
14. Fisher, R. A. *The Genetical Theory of Natural Selection* 2nd edn (New York, Dover, 1958).
15. Lande, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3721-3725 (1981).
16. Kirkpatrick, M. *Evolution* **36**, 1-12 (1982).
17. Zahavi, A. *J. theor. Biol.* **53**, 205-214 (1975).
18. Andersson, M. B. *Evolution* **40**, 804-816 (1986).
19. West-Eberhard, M. J. *Proc. Am. phil. Soc.* **123**, 222-234 (1979).
20. Endler, J. A. & McClellan, T. A. *Rev. Ecol. Syst.* **19**, 395-421 (1988).
21. Arak, A. in *Mate Choice* (ed. Bateson, P.) 181-210 (Cambridge University Press, 1983).
22. Ryan, M. J. *Oxford Surv. Evol. Biol.* (in the press).
23. Cannatella, D. C. & Duellman, W. E. *Copeia* **1984**, 902-921 (1984).
24. Ryan, M. J. & Drewes, R. C. *Biol. J. Linn. Soc. Lond.* (in the press).

ACKNOWLEDGEMENTS. We thank the governments of Panama and Ecuador for permission to conduct these studies, and J. Bull, R. Crocroft, M. Kirkpatrick and S. Weller for comments on the manuscript. The work was supported by the NSF and the Smithsonian Scholarly Studies Program.

Functions of the colour-opponent and broad-band channels of the visual system

Peter H. Schiller, Nikos K. Logothetis
& Eliot R. Charles

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

THE colour-opponent and broad-band channels of the primate visual system originate in the retina and remain segregated through several neural stations in the visual system¹⁻⁷. Until now inferences about their function in vision have been based primarily on studies examining single-cell receptive field properties which have shown that the colour-opponent retinal ganglion cells have small receptive fields, produce sustained responses and receive spatially segregated inputs from different cone types; the broad-band cells have large receptive fields, respond transiently and receive cone inputs that are not spatially separated⁸⁻¹¹. We have now examined the visual capacities of rhesus monkeys before and after interrupting either of these channels with small lesions at the lateral geniculate nucleus. Here we report that the colour-opponent channel is essential for the processing of colour, texture, fine pattern and fine stereopsis, whereas the broad-band channel is crucial for the perception of fast flicker and motion. Little or no deficits were found in brightness and coarse-shape discrimination, low spatial frequency stereopsis and contrast sensitivity after the disruption of either of the channels.

Seven rhesus monkeys were trained to perform visual detection or discrimination tasks. The animals were seated in a primate chair so that they faced either a colour monitor or a panel containing an array of red-green light-emitting diodes (LEDs). Each trial was initiated by the appearance of a small central spot. After the fixation of the spot, as assessed by measuring eye movements using the scleral search-coil technique, stimuli were presented in various parts of the visual field. A single target stimulus (0.25–1.8-degree size) appeared for the detection task. For the discrimination task, eight equally spaced stimuli were presented equidistant from the fixation spot, one of which—the target—was different from the other seven identical non-target stimuli. A monkey making a saccadic movement directly towards the target was rewarded with a drop of apple

juice. The making of other saccadic movements signalled the end of the trial and an error was recorded. The experiment was controlled with two PDP11/73 computer systems.

The animals were trained for several months during which normative data were collected showing that performance in various regions of the visual field at isoeccentric points was similar. Small lesions were then made in either the parvocellular or the magnocellular layers of the lateral geniculate nucleus, which respectively relay the colour-opponent and broad-band channels to the cortex¹²⁻¹⁶. While the animals were anaesthetized, recordings with a micropipette were made to select the appropriate lateral geniculate nucleus layers and visual-field representations to be lesioned. Small quantities (1–3 μ l) of the neurotoxin ibotenic acid were then injected using the same micropipette. Examples of a parvocellular and magnocellular lesion made this way in the left and right lateral geniculate nuclei of one animal are shown in Fig. 1a, b. In most of the animals one lesion was made in each lateral geniculate nucleus in separate sessions.

After each lesion had been made, monkeys were tested for several months with 1,000–4,000 trials taken daily. First the affected area of the visual field was mapped out by systematic exploration of the visual field with small stimuli sensitive to the deficits. Such a map is shown in Fig. 1c for an animal whose lesion is illustrated (Fig. 1a, b). The parvocellular lesions affected visual-field areas ranging from 0.5–9 degrees of eccentricity, and the magnocellular lesions affected visual-field areas from 3 to 15 degrees of eccentricity. Most of the testing was done between 2.5 and 6.5 degrees of eccentricity. After mapping out the affected area, the responses to stimuli presented concurrently to the 'intact' and the 'lesioned' portions of the visual field at isoeccentric locations were recorded. Post-operative performance for stimuli at intact sites was similar to pre-operative performance indicating that at normal sites no significant alterations occurred in the animals's performance as a result of the surgical procedures. We tested animals on the following functions: luminance contrast sensitivity, flicker detection, brightness, colour, pattern, shape and texture discrimination, motion, and stereopsis.

For tests of luminance contrast sensitivity, the animal had to detect a small checkerboard pattern of varied contrasts and spatial frequencies which appeared on a homogeneous background of equal average luminance (38–44 cd m^{-2}). Guessing was minimized by the use of catch trials during which only the fixation spot appeared—monkeys that maintained fixation were

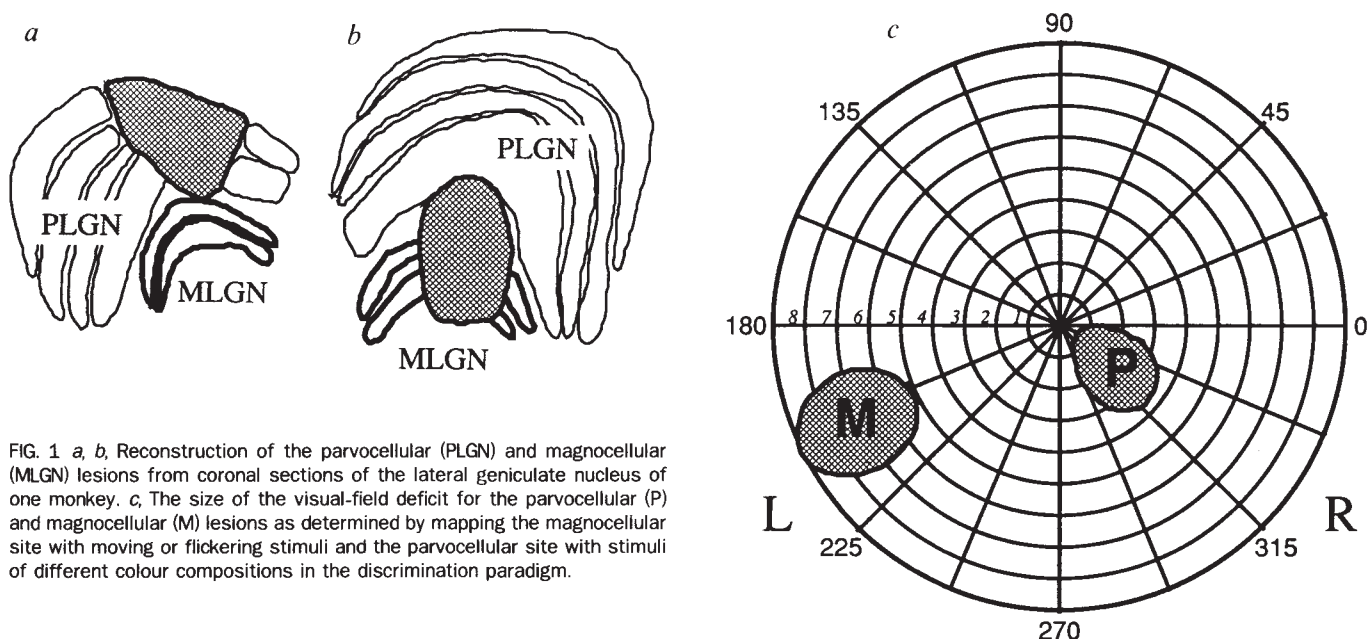


FIG. 1 a, b, Reconstruction of the parvocellular (PLGN) and magnocellular (MLGN) lesions from coronal sections of the lateral geniculate nucleus of one monkey. c, The size of the visual-field deficit for the parvocellular (P) and magnocellular (M) lesions as determined by mapping the magnocellular site with moving or flickering stimuli and the parvocellular site with stimuli of different colour compositions in the discrimination paradigm.

rewarded. Figure 2 shows one set of luminance contrast sensitivity functions: we found a significant loss of contrast sensitivity for stimuli presented in the parvocellular lesion site, especially at high spatial frequencies, confirming an earlier report¹⁷; magnocellular lesions produced no discernible deficits. These observations are surprising because single cells within the broad-band system (which forms 8–10% of the cells in the geniculostriate pathway) are more sensitive than colour-opponent cells^{18–20}. Our findings are supported by a recent report²¹.

Flicker-detection tests used an array of 64 red-green LEDs. For monochromatic flicker the LEDs were set to green. After onset and fixation of a central red LED, one of the green LEDs began to flicker on and off at one of several preset rates and detection was indicated by making a saccadic movement towards it. Data for performance for various flicker rates are shown in Fig. 2: magnocellular lesions dramatically reduced critical flicker fusion (CFF) from the normal 21–23 Hz to 6 Hz (50% detectability point). Parvocellular lesions produced no deficit. Tests of heterochromatic red-green flicker were also made. The intensities of the red and green lights were near isoluminance and all the stimuli were set to yellow by turning on both the red and green portions of the LEDs. After fixation, one of the stimuli began to flicker between red and green. This heterochromatic flicker yielded CFFs (50% threshold) of 10–12 Hz when presented in the normal and parvocellular lesioned portions of the visual field and CFFs of less than 5 Hz when presented in the magnocellular lesioned areas. Magnocellular lesions produced similar results for all cases. No deficits were observed after parvocellular lesions in any animal. These findings are in agreement with Lee *et al.*²² who, on the basis of single-unit recordings in the ganglion cell layer of the macaque retina, propose that the broad-band system underlies performance in the psychophysical task of heterochromatic flicker photometry.

For the discrimination of brightness, an array of eight stimuli appeared which were equidistant from fixation, one of which was more intense than the others. The stimuli were solid squares ranging in size from 0.25 to 1 degree and appeared on a background of 6–9 cd m^{-2} , with the target stimulus set 68–72 cd m^{-2} . The luminance of the non-target stimuli systematically varied, randomized by trial. For colour discrimination, isoluminant stimuli of different wavelength compositions were used (for example, one yellow stimulus and seven green or red or blue stimuli). For pattern discrimination, high contrast (83%) checkerboard patterns were presented, one of which contained

a different spatial frequency checkerboard from the others. We tested stimuli ranging in frequency from 1.0 to 4.6 cycles per degree. For the study of shape perception the animals had to discriminate solid forms (for example, squares versus circles, and squares versus triangles). The height and width of these shapes ranged between 0.5 and 1.8 degrees of visual angle. The different shapes had the same area and appeared at a high contrast (68–72 cd m^{-2} on 6–9 cd m^{-2} background). In all cases the position of the target stimulus was randomized on successive trials, thus appearing in either the intact or lesioned portions of the visual field at isoeccentric points. For the discrimination of texture differences an array of diagonal lines filled the colour monitor screen (///). In one small part of this field the lines appeared with the opposite slant (\\\). Animals were tested with lines of various lengths (16.5 and 36.3 minutes of visual angle), widths (3.3 and 13.2 minutes), densities (number of line segments per degree), colours, and contrasts.

Brightness discrimination data for one animal are shown in Fig. 2. Neither lesion produced deficits in brightness perception indicating that both the colour-opponent and broad-band channels can support the perception of brightness. The data for colour, texture and pattern discrimination shown in Fig. 3 for another animal show what was found throughout: parvocellular lesions devastate performance on all of these tasks; magnocellular lesions do not produce any deficits. Extensive testing using small changes in colour hues, fine pattern differences and high spatial frequency textures also failed to uncover deficits at the magnocellular lesion sites. The striking parvocellular lesion deficits in colour and pattern discrimination seem to be considerably more pronounced than those reported for lesions of cortical area V4 through which the colour-opponent information is believed to be conveyed to the temporal lobe^{23,24}. The deficits in shape perception were also limited to the parvocellular lesioned sites but were observed only for the discrimination of squares from circles (a task that was more difficult, as made evident by the lower performance levels and longer saccadic latencies), suggesting that the broad-band system is capable of discriminating coarse but not fine shapes.

To test motion detection, a random array of spots filled the screen after fixation. In one small region the spots moved coherently. The size of the spots (3.3–13.2 minutes of visual angle in diameter), their frequency, colour, contrast and velocity of movement could be varied. Detection was indicated by a direct saccadic movement towards this location. The results in Fig. 3

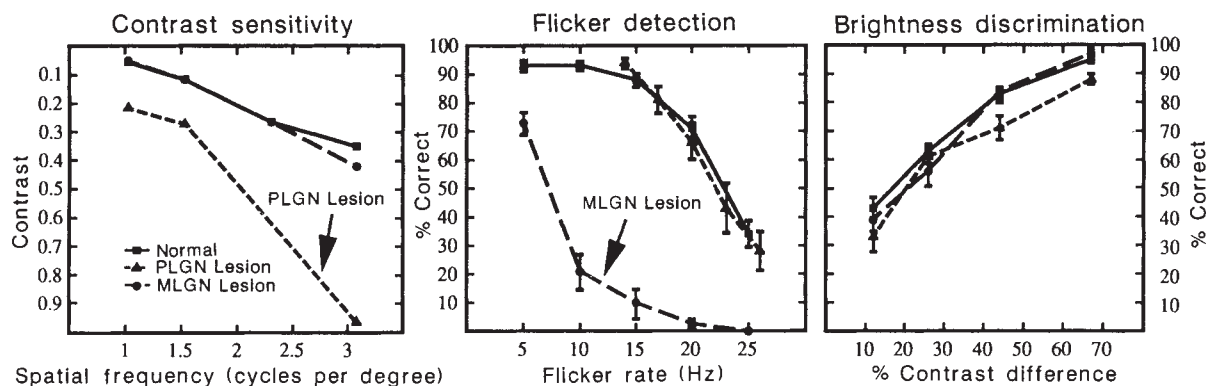


FIG. 2 Contrast sensitivity, flicker detection and brightness discrimination functions derived from the normal, parvocellular and magnocellular lesioned visual-field sites of one monkey. Contrast sensitivity functions show contrast level for absolute threshold (50% detectability), where contrast is equivalent to $(\text{luminance 1} - \text{luminance 2}) / (\text{luminance 1} + \text{luminance 2})$ of the checkers in the checkerboard patterns used in the detection task ($N=6,000$). For flicker detection and brightness discrimination, each of the data points shown is based on a minimum of 100 trials collected in blocks of 10 trials per

condition presented in random order (for 4–8 positions and 1–8 stimulus conditions). The error bars show ± 1 s.e.m. The target stimulus values for brightness discrimination are expressed as contrast differences compared with the non-targets as shown on the abscissa. The luminance of the background for contrast sensitivity was 38–44 cd m^{-2} and 6–9 cd m^{-2} for the other tasks. Target luminance for brightness discrimination was 68–72 cd m^{-2} . Non-target luminance values, as shown on abscissa, are expressed as a percentage of contrast relative to target luminance.

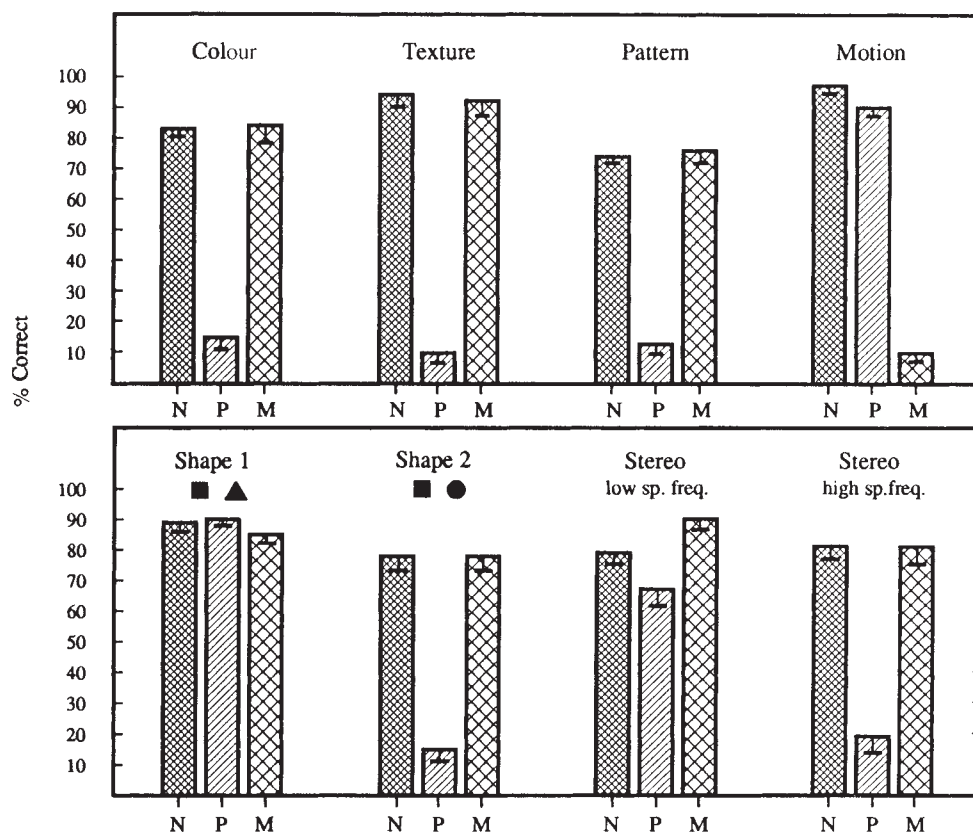


FIG. 3 Colour, texture, pattern, motion, shape and stereoscopic-depth-perception performance derived from the normal (N), the parvocellular lesioned (P), and magnocellular lesioned (M) representation of the visual field of one monkey. For colour discrimination isoluminant green-yellow discrimination is shown. Other colour pairings yielded similar deficits. For pattern discrimination the spatial frequency of the target and non-target stimuli was 1 and 4.6 cycles per degree. Other spatial frequency (sp. freq.) pairings produced similar deficits. Motion detection is for movement of high contrast stimuli at 6.5 degrees per second of (83% contrast). The individual random-dot size was 13.2 minutes for low- and 6.6 minutes for high spatial frequency stereo. Disparity was 6.6 minutes. Contrast was 83% for pattern, texture, motion, shape and stereo. Background was set at 6–9 cd m^{-2} . $N \geq 100$ per bar.

show data for the detection of motion of 6.5 degrees per second. Pronounced deficits in motion detection were always limited to stimuli presented in magnocellular lesioned sites; no deficits were ever produced as a result of parvocellular lesions. Motion discrimination was also tested by moving one of eight stimuli either at a different velocity or in a different direction from that of the others. Deficits in motion discrimination were also found for stimuli presented in magnocellular lesioned sites.

To test stereopsis, pairs of random dot stereograms were generated on the colour monitor, which the monkey viewed through a stereoscope. Within a small square area the dots were horizontally displaced so as to appear in depth. Monkeys making a saccadic movement towards this region were rewarded. The colour of the stimuli and their spatial frequency were systematically varied. Disparities from 6.6 to 39.6 minutes were tested and spatial frequencies for the random-dot stereograms ranged between 2.2 and 4.5 cycles per degree. The summary of data in Fig. 3 shows that parvocellular lesions produced severe deficits in stereopsis with high but not with low spatial frequency stereograms. None of the magnocellular lesions yielded any deficits in stereopsis at the spatial frequencies and disparities tested. A discrimination paradigm was also used to study stereopsis. Eight stimuli were presented in depth, one of which—the target—had a different disparity and hence appeared at a different depth. This task also revealed deficits in stereopsis for stimuli presented in parvocellular lesioned sites, but showed no loss in performance at magnocellular lesioned sites when compared with the performance at isoeccentric intact locations. These results were consistent among the animals, and do not support the claim that the broad-band system is important for static stereopsis².

In two control animals we made combined parvocellular and magnocellular lesions with a radio frequency heat probe. These animals were unable to either detect or discriminate in the scotoma created in this way any of the stimuli used in this study.

The results described above show that the perception of colour, texture, shape, pattern and fine stereopsis is mediated by the colour-opponent channel. The temporal aspects of perception (motion and flicker) are mediated predominantly by the broad-band channel. At low spatial frequencies both channels are capable of processing stereoscopic depth, brightness and scotopic vision. Although our results do not provide direct evidence about the manner in which the information carried in these two channels accesses higher neural centers, they indicate that if some cortical areas are indeed specialized for the analysis of shape, brightness and stereoscopic depth², they are likely to receive convergent input from the color-opponent and broad-band channels. □

Received 16 August; accepted 19 October 1989.

1. Rodieck, R. W. A. *Rev. Neurosci.* **2**, 193–225 (1979).
2. Livingstone, M. & Hubel, D. *Science* **240**, 740–749 (1988).
3. Martin, K. A. C. *Trends Neurosci.* **11**, 380–387 (1988).
4. Lennie, P. *Vision Res.* **20**, 561–564 (1980).
5. DeYoe, E. A. & Van Essen, D. C. *Trends Neurosci.* **11**, 219–226 (1988).
6. Ungerleider, L. G. & Mishkin, M. in *Analysis of Visual Behavior* (eds Ingle, D. J., Goodale, M. A. & Mansfield, R. J. W.) (Massachusetts Institute of Technology, Cambridge, Massachusetts, 1982).
7. Zeki, S. M. & Shipp, S. *Nature* **335**, 311–317 (1988).
8. De Monasterio, F. M. & Gouras, P. J. *J. Physiol., Lond.* **251**, 167–195 (1975).
9. Gouras, P. J. *J. Physiol., Lond.* **204**, 407–419 (1969).
10. Schiller, P. H., Malpeli, J. G. & Schein, S. J. *J. Neurophysiol.* **42**, 1124–1133 (1977).
11. Derrington, A. M., Krauskopf, J. & Lennie, P. *J. Physiol., Lond.* **357**, 241–265 (1984).
12. Dreher, B., Fukuda, Y. & Rodieck, R. W. *J. Physiol., Lond.* **258**, 433–452 (1976).
13. Leventhal, A. G., Rodieck, R. W. & Dreher, B. *Science* **213**, 1139–1142 (1981).
14. Schiller, P. H. & Malpeli, J. G. *J. Neurophysiol.* **40**, 428–445 (1977).
15. Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **29**, 1115–1156 (1966).
16. Hubel, D. H. & Wiesel, T. N. *J. comp. Neurol.* **146**, 421–450 (1972).
17. Merigan, W. H. & Eskin, T. A. *Vision Res.* **26**, 1751–1761 (1986).
18. Shapley, R., Kaplan, E. & Soodak, R. *Nature* **292**, 543–545 (1981).
19. Schiller, P. H. & Colby, C. L. *Vision Res.* **23**, 1631–1641 (1983).
20. Derrington, A. M. & Lennie, P. *J. Physiol., Lond.* **357**, 219–240 (1984).
21. Merigan, W. H., Katz, L. M. & Maunsell, J. H. R. *Invest. Ophthalmol. Vis. Sci. Suppl.* **30**, 53 (1989).
22. Lee, B. B., Martin, P. R. & Valberg, A. *J. Physiol., Lond.* **404**, 323–347 (1988).
23. Heywood, C. A. & Cowey, A. *J. Neurosci.* **7**, 2601–2617 (1987).
24. Schiller, P. H., Charles, E. R. & Logothetis, N. K. *Invest. Ophthalmol. Vis. Sci. Suppl.* **29**, 328 (1988).

Defective processing and presentation of exogenous antigens in mutants with normal HLA class II genes

Elizabeth Mellins*, Laura Smith*, Benjamin Arp*, Tom Cotner*, Esteban Celis† & Donald Pious*†§

Departments of Pediatrics*, Genetics† and Immunology§, University of Washington, Seattle, Washington 98195, USA
‡ T-Cell Sciences, Cambridge, Massachusetts 02139, USA

PRESENTATION of an exogenous protein antigen to helper (CD4⁺) T-lymphocytes by antigen presenting cells (APC) generally requires that the APCs degrade the native protein antigen into an immunogenic peptide, a process termed 'antigen processing', and that this peptide bind to a major histocompatibility complex (MHC) class II molecule¹⁻³. The complex of peptide and MHC molecule on the APC surface provides the stimulatory ligand for the $\alpha\beta$ T cell receptor⁴. The intracellular pathways and molecular mechanisms involved in the generation of the peptide-MHC complex are not well understood. Here, we describe several mutant APCs which are altered in their ability to present native exogenous protein antigens but effectively present immunogenic peptides derived from these proteins. The lesions in these mutants are not in the class II structural genes, but they affect the conformation of mature class II dimers.

In the course of selecting HLA-DR3 β 1 structural gene mutants from a HLA-DR3-expressing B lymphoblastoid cell

line (B-LCL), 8.1.6, we discovered a class of mutants with unexpected properties, the main one being a dramatic defect in the ability to present native protein antigens to HLA class II-restricted T cells. The defects in antigen presentation in the mutants are not caused by a reduction in the cell-surface levels of class II molecules, as these levels are normal (Table 1 and ref. 5). We have isolated eight independent mutants with this 'presentation-defective' phenotype by replicate selections: young clones of cell line 8.1.6 were mutagenized with ethyl methane sulphonate and immunoselected with monoclonal antibody 16.23 plus complement (see ref. 5 for details). Monoclonal antibody 16.23 recognizes a polymorphic determinant on DR3 β 1 molecules and some DR3 β specificities, but not on DQ or DP molecules⁶. Although the mutants were selected with an antibody which recognizes only DR β chains, the defect in this subset of mutants affects not only DR-, but also DQ- and DP-restricted antigen presentation. Fig. 1a shows a comparison of T-cell stimulation by 9.5.3, a representative of this class of mutant, and its progenitor 8.1.6. The 9.5.3 mutant is unable to present protein antigens to DR3-, DQ2-, or DP4-restricted T-cell clones specific for a variety of protein antigens (see legend to Fig. 1). The functional defect in these mutants is profound but not absolute. When antigen-specific, class II-restricted polyclonal T-cell lines are used as responders (see TT/class II, Fig. 1a), the mutants generally demonstrate some APC activity (0-35% of 8.1.6 activity, depending on the T-cell line). No antigen-specific clones responsive to these mutant APCs have been isolated thus far from these T-cell lines.

The functional defects in these mutants and the apparent class II conformational changes (see below) affect several class II molecules. The protocol used for generating the mutants made it seem unlikely that each harbours structural gene mutations in several class II genes, however^{5,7}. Indeed, we sequenced the

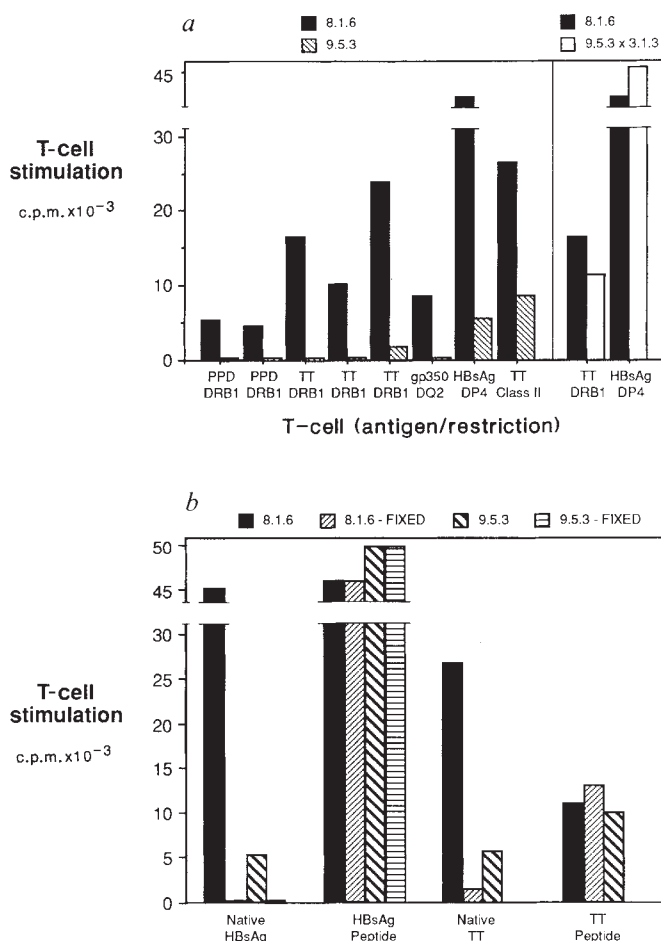
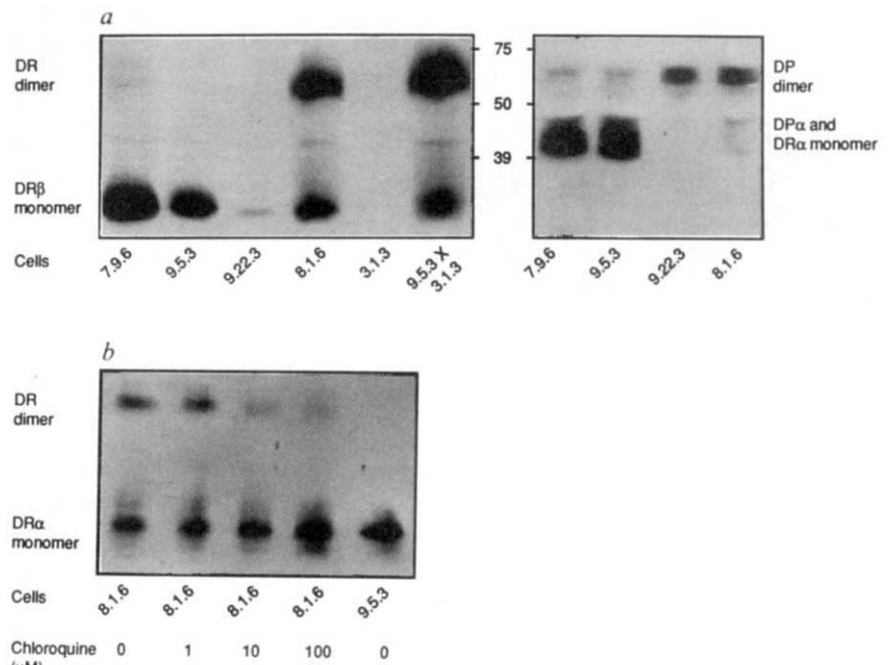


FIG. 1 HLA class II-restricted presentation of intact protein antigens and immunogenic peptides by mutants and a mutant derived somatic cell hybrid. **a**, Proliferation of T cells stimulated by mutant 9.5.3, its progenitor 8.1.6, or hybrid 9.5.3 x 3.1.3 as APC. T cells are designated by antigen and restricting element and include seven T-cell clones [specific for purified protein derivative of *M. tuberculosis* (PPD), tetanus toxoid (TT), a major envelope glycoprotein of EB virus (gp 350), and hepatitis B surface antigen (HBsAg)] and a polyclonal, tetanus toxoid specific T-cell line, TT/class II. Clones with the same apparent specificity were isolated independently. T-cell stimulation was measured by [³H]thymidine incorporation (c.p.m. in the presence of soluble antigen - c.p.m. in the absence of soluble antigen). Incorporation in the absence of antigen was approximately equal to background c.p.m. from APC alone and was generally <500 c.p.m. Data are median values from triplicate cultures from a representative experiment; s.e.m. of relative responses (response to mutant/response to 8.1.6) were <15% in three or more replicate experiments; **b**, Proliferation of antigen specific T cells stimulated with native antigen or immunogenic peptide, presented by progenitor 8.1.6, glutaraldehyde-fixed 8.1.6, mutant 9.5.3 and glutaraldehyde-fixed 9.5.3. T-cell stimulation was calculated as in (a). **METHODS.** **a**, The progenitor line 8.1.6 is hemizygous for DR and DQ A and B genes and expresses the HLA class II specificities DR3, DQ2 and DP4.1 (ref. 18). Isolation of mutant 9.5.3 is described in the text. With the exception of gp350/DQ2, the antigen-specific clones were isolated as described¹⁹, and their restricting elements were identified using HLA class II molecule-loss mutants²⁰. T-cell clone gp350/DQ2 (manuscript in preparation) was provided by B. Carreno. The polyclonal class II-restricted tetanus toxoid-specific T-cell line was isolated as described²⁰. For assay, 2 x 10⁴ T-cell blasts were co-cultured with 10⁵ mitomycin-C treated B-LCLs as APC in the presence or absence of soluble antigen. **b**, To identify the peptide (amino acid position 14-33) that stimulates a HBsAg specific, DP4-restricted T-cell clone, we screened synthetic oligopeptides corresponding to short fragments (10-30 amino acids) of HBsAg²¹. The peptide corresponding to residues 1,274-1,285 of tetanus toxoid in association with DR3 β 3 stimulates the polyclonal, DR restricted TT-specific T-cell line shown (ref. 8 and our unpublished results). Peptides were synthesized by the solid-phase method. In proliferation assays, recombinant HBsAg, the gift of Merck, Sharp & Dohme, was used at 10 μ g ml⁻¹, HBsAg peptide was used at 1 μ g ml⁻¹ and TT peptide was used at 20 μ g ml⁻¹. Glutaraldehyde fixation of cells was performed as described²². T-cell proliferation was assayed as in **a**.

FIG. 2 HLA-DR and DP dimers extracted from mutants and chloroquine-treated progenitor 8.1.6 dissociate into monomeric subunits in 1-dimensional SDS-PAGE. *a*, Western blot analysis of detergent (NP-40) extracts of the indicated cell lines. Mutants 7.9.6 and 9.5.3, independently derived presentation defective mutants; 9.22.3, 8.1.6-derived, DR-null mutant (ref. 5); 8.1.6, progenitor line; 3.1.3, DR null mutant (fusion partner); 9.5.3 $\times$ 3.1.3, hybrid. Unboiled, whole-cell extracts were analysed by 1-dimensional SDS-PAGE and immunoblotted with monoclonal antibody HB10.A (left) which reacts with DR dimers and DR β monomers²³ or monoclonal antibody B7/21.2 (right) which reacts with DP dimers and DP α and DR α monomers (ref. 24 and our unpublished results). The stability of HLA class II molecules to SDS without boiling has been described^{25,26}. The differences in intensities of the DR β monomer bands in the two mutants and of the DR dimer bands of 8.1.6 and the hybrid probably reflect small differences in loading, as they were not seen consistently. The low level of detectable DR β monomer in the homozygous DR α deletion mutant 9.22.3 reflects instability of the unpaired DR β chain (ref. 5). Because B7/21.2 recognizes DR α as well as DP α monomers, the altered stability of the DP molecule in the mutants is best evaluated by the disappearance of DP dimer. Relative molecular masses, K, are shown. *b*, Western blot analysis of detergent extracts of 8.1.6 cells, after 48 h incubation with increasing concentrations of chloroquine. The blot was probed with monoclonal antibody DA6.147, which recognizes DR dimer and DR α monomer²⁷. The apparent difference in the relative amounts of DR dimer and monomer in progenitor 8.1.6 in *a* and *b* reflects the higher affinity of DA6.147 for DR α in monomeric than dimeric form, as compared to the more equal affinities of HB10.A for DR β in dimeric and monomeric forms, (ref. 27 and our unpublished observations). The effect on dimer stability observed with chloroquine treatment is specific; no change in the relative amounts of dimeric and monomeric DR is observed in 8.1.6 cells incubated for 48 h with mitomycin C, cyclohexamide, or tunicamycin over a range of concentrations chosen to produce a 0–50%



drop in cell viability (data not shown). The cell-surface binding of anti-DR specific monoclonal antibody, VI.15, to 8.1.6 treated with 5×10^{-4} M chloroquine, expressed as a proportion of binding to untreated 8.1.6, is $86\% \pm 4$.

METHODS. Western blot analysis were performed essentially as described⁵. The generation of somatic cell hybrid 9.5.3 $\times$ 3.1.3, is described in the legend to Fig. 3. For chloroquine treatment, 8.1.6 cells were incubated for 24 h in RPMI 1640 with 15% FCS and the indicated amounts of chloroquine. Cells were then pelleted and resuspended in fresh medium plus chloroquine for an additional 24 h. Under these conditions, cell viabilities were $>90\%$ at 48 h at all doses of chloroquine assayed.

entire coding regions of the *DRA* and *DRB* genes of two independent mutants (9.5.3 and 7.9.6) and found no mutations (data not shown). These findings, together with the functional defects, indicate that the mutations disrupt a process required for the presentation of native protein antigens. We asked therefore whether the inability of the mutants to stimulate antigen specific T-cell clones could be overcome by substituting an immunogenic peptide for native protein antigen. We used a synthetic peptide corresponding to amino acids 14–33 of hepatitis B surface antigen (HBsAg), which stimulates a DP4-restricted T-cell clone. We first determined that the native antigen requires processing by APC for presentation. Glutaraldehyde fixation of the progenitor line 8.1.6 ablates its ability to present

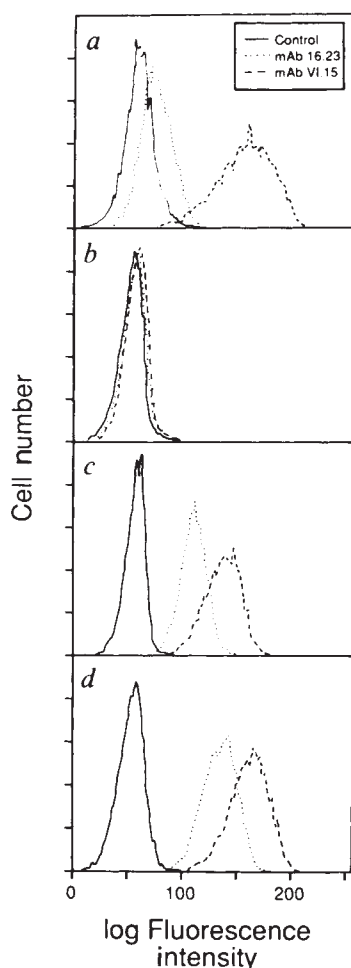
native HBsAg to specific T cells, but fixed 8.1.6 cells readily present the immunogenic peptide from HBsAg (Fig. 1b). The mutants, like fixed cells, are unable to present the native antigen, but are able to present the peptide antigen as effectively as 8.1.6. (Fig. 1b; data shown are from a representative mutant, 9.5.3). A second synthetic peptide, corresponding to amino acids 1,274–1,285 of tetanus toxoid antigen and presented in association with HLA-DR molecules⁸ is also presented normally by mutant 9.5.3 (Fig. 1b). Thus, the class II molecules expressed by the mutants can function normally in antigen presentation if provided with exogenous immunogenic peptides. These data, taken together with those from Fig. 1a, indicate that the presentation defects in the mutants derive from the inability to generate MHC class II-peptide complexes from native protein antigens.

Although the *HLA-DR* structural genes are normal in the mutants and the level of cell-surface DR expression is normal (Table 1), two lines of evidence indicate that the mature class II molecules in these cells differ from those expressed by progenitor 8.1.6. First, the HLA-DR dimers expressed by the mutants have lost expression of two polymorphic DR3 determinants, recognized by monoclonal antibodies 16.23 and 7.3.19, respectively (Table 1). The mutants maintain expression of other polymorphic epitopes, as demonstrated by normal binding of the DR3 specific monoclonal antibody CD6.B1 (Table 1) and 12 anti-DR3 alloantisera (data not shown). Second, DR and DP dimers extracted from the mutants dissociate into monomers in SDS-PAGE, whereas the majority of the DR and DP molecules in the progenitor line retain the dimeric state under these conditions (Fig. 2a and ref. 5). Given that the primary structure of the DR molecules is normal in these mutants, the alterations in antibody binding and in the stability of extracted

TABLE 1 Cell-surface antibody-binding to mutant cells

Cell line	Antibody-binding ratio ($\times 100$)			
	16.23	7.3.19	CD6B1	VI.15
8.1.6	(100)	(100)	(100)	(100)
9.5.3	0 ± 5	22 ± 2	105 ± 5	102 ± 5

Binding ratio: (c.p.m. bound by mutant – c.p.m. bound by negative control)/(c.p.m. bound by 8.1.6 – c.p.m. bound by negative control). The negative control is 9.22.3, a DR null mutant derived from 8.1.6 (ref. 5). The binding ratios for 9.5.3 represent median values $\pm$ s.e.m. from three or more independent assays. For cell lines 8.1.6 and 9.5.3, cell-surface radioimmune binding was performed as described¹⁵. Monoclonal antibodies 16.23 (ref. 6), 7.3.19 (ref. 16), and CD6B1 (ref. 17) react with polymorphic, DR β determinants; their specificities are distinct, as evidenced by their inability to cross-inhibit each other (unpublished data). VI.15 (ref. 15) reacts with a monomorphic DR determinant, expressed on DR dimers only.



dimers strongly indicate that the lesions in the mutants affect the conformation of mature HLA class II molecules. This is not a drastic effect, however, as evidenced by the normal cell-surface binding of some conformation-sensitive antibodies (for example, VI.15), and the ability of the mutants to present peptide antigen normally.

In all eight independent mutants, the defect in presentation of soluble proteins is associated with an altered conformation of class II molecules. One attractive hypothesis that could account for this association is that impaired generation of class II-peptide complexes from native proteins (including 'self pro-

teins') results in altered occupancy of the peptide-binding sites and associated conformational changes in class II molecules. This is consistent with data suggesting that the peptide binding groove of purified MHC molecules is normally occupied^{9,10}. We attempted to induce a similar conformational change in class II molecules by treating the progenitor line, 8.1.6, with chloroquine, an agent known to disrupt antigen processing for class II restricted presentation¹¹. To examine cells in which the majority of DR molecules were synthesized in the presence of chloroquine, we propagated 8.1.6 cells for 48 h in medium containing chloroquine over a range of concentrations. Western blots of DR molecules from detergent extracts of 8.1.6 cells grown in 1×10^{-5} M chloroquine resemble those of the mutants (Fig. 2b). These blots demonstrate a loss of DR dimer and a compensatory increase in DR α monomer, indicating an alteration in stability of the extracted class II molecules in SDS. Thus, interrupting the formation of MHC-peptide complexes by both mutation and chloroquine is associated with altered stability of extracted class II dimers. Because chloroquine has also been reported to delay dissociation of invariant chain from class II dimers¹², however, the mechanisms inducing dimer instability in the two cases may not be identical.

METHODS. 3.1.3 is a DR null mutant cell line derived from 3.1.0; 3.1.0 has a deletion of one complete HLA haplotype¹⁸. Mutant 3.1.0 expresses the specificities DR1, DQ1, DP4.1. Mutant 3.1.3 expresses no DR β mRNA or protein (not shown). For somatic cell hybridization, dominant selectable markers were introduced into 9.5.3 and 3.1.3. An *Escherichia coli* gpt gene, conferring resistance to mycophenolic acid, was transfected into 9.5.3 by electroporation as described (Stimac, E., Urieli-Shoval, S., Kempin, S. A. & Pious, D. A., in preparation). The dihydrofolate reductase gene, conferring resistance to methotrexate, was introduced into 3.1.3 by coculture with PA317-SDHT, a retrovirus packaging fibroblast line, provided by A. Dusty Miller²⁹. Somatic cell hybrids were prepared by suspension cell fusion³⁰, using PEG 1500 for 90 s at 37 °C. For antibody binding, cells were incubated with saturating amounts of FITC-conjugated antibody, washed and analysed on a Becton-Dickinson FACSTAR. Cell number is displayed against a 4 log unit axis of fluorescence intensity.

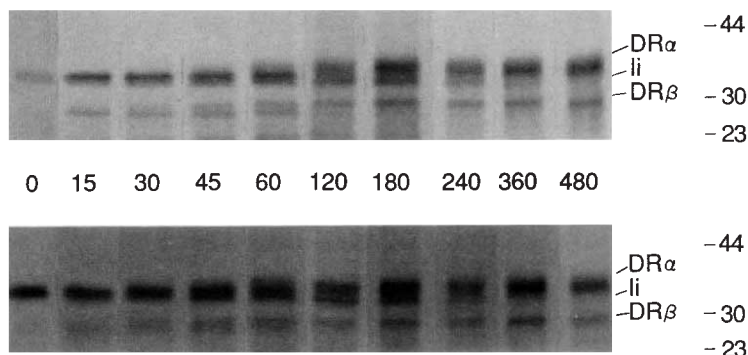
teins') results in altered occupancy of the peptide-binding sites and associated conformational changes in class II molecules. This is consistent with data suggesting that the peptide binding groove of purified MHC molecules is normally occupied^{9,10}. We attempted to induce a similar conformational change in class II molecules by treating the progenitor line, 8.1.6, with chloroquine, an agent known to disrupt antigen processing for class II restricted presentation¹¹. To examine cells in which the majority of DR molecules were synthesized in the presence of chloroquine, we propagated 8.1.6 cells for 48 h in medium containing chloroquine over a range of concentrations. Western blots of DR molecules from detergent extracts of 8.1.6 cells grown in 1×10^{-5} M chloroquine resemble those of the mutants (Fig. 2b). These blots demonstrate a loss of DR dimer and a compensatory increase in DR α monomer, indicating an alteration in stability of the extracted class II molecules in SDS. Thus, interrupting the formation of MHC-peptide complexes by both mutation and chloroquine is associated with altered stability of extracted class II dimers. Because chloroquine has also been reported to delay dissociation of invariant chain from class II dimers¹², however, the mechanisms inducing dimer instability in the two cases may not be identical.

FIG. 4 The kinetics of invariant chain (Ii) association and dissociation from HLA-DR dimers are the same in mutant 9.5.3 and progenitor, 8.1.6. Mutant and progenitor cells were pulse-labelled with [³⁵S]-methionine for 10 min and HLA-DR complexes were immunoprecipitated with mAb VI.15 after the indicated periods of chase. Immunoprecipitated material was analysed by 1-D SDS-PAGE; the portions of the gels in the relevant molecular weight range for DR α and DR β monomers are shown. The designations of bands at the appropriate apparent molecular mass as DR α (~35,000), DR β (~29,000) and Ii (~33,000) were confirmed by western blot analysis of VI.15 immunoprecipitates from unlabelled cells (not shown). For western blots, monoclonal antibodies DA6.147, anti-DR α (ref. 25), HB10.A, anti-DR β (ref. 23) and B73-BU-43 and B113-BU-45, anti-invariant chain antibodies³¹ were used. At $t=0$, the DR β band from both cell lines is faint, but detectable. The appearance, over time, of higher apparent molecular weight species of DR α and DR β reflects sialylation of the polypeptide chains.

9.5.3
(mutant)

Time (min)

8.1.6
(progenitor)



METHODS. Cells were radiolabelled with [³⁵S]-methionine (0.5 m Ci per 18×10^6 cells) and chased with RPMI 1640 medium containing unlabelled methionine at 3.5 times the concentration of methionine in standard RPMI. Immunoprecipitation of HLA-DR antigens was performed as described³². Samples were analysed on a 10% SDS-PAGE gel.

To determine whether the mutation in a presentation defective mutant is dominant or recessive, we generated a somatic cell hybrid between the mutant 9.5.3 and a DR null mutant derived from a DR1 expressing (16.23 negative) progenitor (3.1.3; Fig. 3 legend), which presents native antigens effectively to DQ- and DP-restricted T cells. The 9.5.3 × 3.1.3 hybrid cell re-expresses the 16.23 epitope which was lost in the 9.5.3 mutant (Fig. 3). The DR molecules expressed by the hybrid do not dissociate in SDS-PAGE (Fig. 2a) and the hybrid presents native protein antigens (Fig. 1a). Rescue of the mutant phenotype by somatic cell hybridization indicates that the mutation in mutant 9.5.3 is recessive. An alternative possibility, that the mutation is dominant but has segregated in the hybrid, is unlikely; hybrids between Epstein-Barr virus-transformed cell lines are karyotypically stable¹³, and we have isolated a second hybrid with the same phenotype from an independent fusion.

The fact that the function and conformation of multiple class II gene products are altered in these mutants raised the possibility that the mutations might have affected the invariant chain (Ii), a nonpolymorphic chain associated with class II antigens during their intracellular transport to the cell surface¹⁴. To evaluate Ii association and dissociation from MHC class II molecules in the mutant cells, we pulse-labelled mutant 9.5.3 and its progenitor 8.1.6 and compared the appearance and disappearance of Ii molecules in immunoprecipitates of HLA-DR molecules. There are no apparent differences in the kinetics of Ii association or dissociation in the two cell lines (Fig. 4). In addition, invariant chains immunoprecipitated from mutants appear normal in quantity and charge in 2-dimensional gels (data not shown), suggesting that the presentation defects in the mutants do not result from mutations in the invariant-chain structural gene.

We have isolated a series of independent mutant APCs which are defective in antigen presentation, although their HLA class II genes are normal. These mutants provide evidence for the involvement of gene products other than the class II molecules themselves in antigen processing and presentation. The presentation defective phenotype of a given mutant could result (1) from a deficiency in generating peptides from native proteins, (2) from loss of access of class II molecules to peptides because of abnormal intracellular trafficking of class II molecules, or (3) from other as yet unknown steps in antigen processing or presentation. The common features of the mutant phenotype permit easy screening for isolation of such mutants. The mutants, in turn, should provide valuable tools to investigate pathways of antigen processing and presentation. □

30. Lane, R., Crissman, R. & Ginns, S. *Meth. Enzym.* **121**, 174-182 (1986).
31. Dorken, B. in *Leucocyte Typing IV* (ed. Knapp, W.) (Oxford University Press, in the press).
32. Cotner, T., Charbonneau, H., Mellins, E. & Pious, D. *J. Biol. Chem.* **264**, 11107-11111 (1989).

ACKNOWLEDGEMENTS. We thank Beatriz Carreno for the gift of the gp350 T cell clone, Barbara Miller for technical assistance, Dan Hill for preparation of the manuscript and Drs Simcha Urieli-Shoval and Kindred Ritchie for help with introduction of dominant selectable markers. This work was supported by research grants and a Physician-Scientist Award (to E.M.) from the National Institute of Health.

Class II MHC molecules can use the endogenous pathway of antigen presentation

Jed G. Nuchtern*, William E. Biddison† & Richard D. Klausner*

* Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

† Neuroimmunology Branch, National Institute of Neurologic Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892, USA

MODELS for antigen presentation have divided the world of antigens into two categories, endogenous and exogenous¹⁻³, presented to T cells by class I and class II major histocompatibility complex (MHC) encoded molecules, respectively. Exogenous antigens are thought to be taken up into peripheral endosomal compartments where they are processed for binding to class II MHC molecules. Endogenous antigens are either synthesized⁴⁻⁶ or efficiently delivered to the cytoplasm^{7,8} before being partially degraded in an as yet undefined way, and complexed with class I MHC molecules. A useful phenotypic distinction between the two pathways has been the sensitivity to weak bases, such as chloroquine, which is a property only of the exogenous pathway¹. The fungal antibiotic brefeldin A (BFA), which blocks protein transport from the endoplasmic reticulum to the Golgi network⁹⁻¹¹, also blocks class I-restricted antigen-presentation, providing us with the corresponding marker of the endogenous pathway^{12,13}. Experiments with influenza virus antigens^{1,14} have supported the view that class II MHC molecules can present exogenous but not endogenous antigen¹⁻³, whereas the observation that class II MHC molecules present measles virus non-membrane antigens by a chloroquine-insensitive pathway¹⁵ suggests that this is not always the case. We show here that influenza A matrix protein can be effectively presented to class II-restricted T cells by two pathways: one of which is chloroquine-sensitive, BFA-insensitive, the other being chloroquine-insensitive and BFA-sensitive. Our results indicate that both class I and class II molecules can complex with antigenic peptides in a pre-Golgi compartment and favour a unified mechanism for MHC-restricted endogenous antigen presentation.

To test whether an endogenous pathway participates in presentation of influenza virus antigens to class II restricted cytotoxic T lymphocytes (CTL), we used as target cells those from the human B lymphoblastoid cell line, K4B, infected with influenza A/JAP/57 virus. On viral infection these cells present two distinct T-cell epitopes corresponding to amino acids 17-31 (refs 16, 17) or 55-73 (refs 18, 19) of the influenza A matrix protein (M1 17-31 and M1 55-73, respectively). Infected cells are susceptible to lysis by M1 55-73 specific cytotoxic T cells that are class I-restricted (HLA-A2) or by M1 17-31 specific cytotoxic T cells that are CD4⁺ and class II (HLA-DR1)-restricted. The cellular and immunological effects of BFA have been characterized (refs 9-13). Although this drug rapidly and reversibly blocks class I-restricted presentation of endogenous antigen, apparently through specific interference with transport of newly synthesized protein from the endoplasmic reticulum (ER) to the Golgi network, it does not block processes that affect viral entry (endocytosis, endosomal acidification, or lysosomal function¹⁰), viral protein synthesis, the ability of peptide-pulsed or allogeneic B lymphoblastoid cells to be lysed as targets,

Received 16 June; accepted 10 November 1989

1. Ziegler, K. & Uhanue, E. R. *Proc. natn Acad. Sci. U.S.A.* **79**, 1869-1875 (1981).
2. Shimonkevitz, R., Kappler, J., Marrack, P. & Grey, H. *J. exp. Med.* **158**, 303-316 (1983).
3. Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E. & Uhanue, E. R. *Nature* **317**, 359-361 (1985).
4. Saito, T., Weiss, A., Miller, J., Norcross, M. A. & Germain, R. N. *Nature* **325**, 125-130 (1987).
5. Pious, D., Dixon, L., Levine, F., Cotner, T. & Johnson, R. *J. exp. Med.* **162**, 1193-1207 (1985).
6. Johnson, J. P., Meo, T., Riethmuller, G., Schendel, D. J. & Wank, R. *J. exp. Med.* **156**, 104-111 (1982).
7. Pious, D., Erlich, H., Gladstone, P. & Levine, F. *Barbury Rep. Cold Spring Harb. Lab.* **14**, 61-68 (1983).
8. Demetz, S. *et al J. Immun.* **142**, 394-402 (1989).
9. Buus, S., Sette, A., Colon, S. M. & Grey, H. M. *Science* **142**, 1045-1047 (1988).
10. Bjorkman, P. J. *et al Nature* **329**, 506-518 (1987).
11. Ziegler, H. K. & Uhanue, E. R. *Proc. natn Acad. Sci. U.S.A.* **79**, 175-178 (1982).
12. Nowell, J. & Quaranta, V. *J. exp. Med.* **162**, 1371-1376 (1985).
13. Gladstone, P. & Pious, D. *Somatic Cell Genet.* **6**, 285-298 (1980).
14. Jones, P. P., Murphy, D. B., Hewgill, D. & McDevitt, H. O. *Molec. Immun.* **16**, 51-60 (1979).
15. Gladstone, P., Furesz, L. & Pious, D. *Proc. natn Acad. Sci. U.S.A.* **79**, 1235-1239 (1982).
16. Koning, F., Schreuder, I., Giphart, M. & Brunning, H. *Hum. Immun.* **9**, 221-230 (1984).
17. Sachs, J. A. *et al Tissue Antigens* **28**, 199-207 (1986).
18. Levine, F. *et al Proc. natn Acad. Sci. U.S.A.* **82**, 3741-3745 (1985).
19. Mellins, E., Arp, B., Ochs, B., Erlich, H. & Pious, D. *J. exp. Med.* **168**, 1531-1537 (1988).
20. Mellins, E., Woelfel, M. & Pious, D. *Hum. Immun.* **18**, 211-223 (1987).
21. Celis, E., Ou, D. & Otvos, L. *J. Immun.* **140**, 1808-1815 (1988).
22. Lanzavecchia, A. *Nature* **314**, 537-539 (1985).
23. Clark, E. A. & Yakoshi, T. in *Leucocyte Typing* (eds Bernard, A., Baunsell, L., Dausset, J. & Schlossman, S.) 195 (Springer, Heidelberg, 1984).
24. Watson, A. J., DeMars, R., Trowbridge, I. S. & Bach, F. H. *Nature* **304**, 358-361 (1983).
25. Billing, R., Safari, M. & Peterson, P. *J. Immun.* **117**, 1589-1593 (1976).
26. Cresswell, P. *Eur. J. Immun.* **7**, 636-639 (1977).
27. Guy, K. *et al Eur. J. Immun.* **12**, 942-948 (1982).
28. Gladstone, P. & Pious, D. *Nature* **271**, 459-461 (1978).
29. Miller, A. D. & Buttimore, C. *Molec. cell. Biol.* **6**, 2895-2902 (1986).

or the ability of cytotoxic T cells to lyse infected or peptide-pulsed targets¹².

Using live influenza virus and K4B cells, we find that BFA significantly inhibits class II-restricted presentation of cytoplasmic viral antigens to M1 17-31 peptide specific DR1-restricted CTL lines (Fig. 1a, b). Addition of this drug decreases cell-mediated killing by 56-93% with an average of 72% in multiple experiments. BFA did not affect CTL function *per se*, as M1 17-31 pulsed cells (Fig. 1a, b) and untreated infected cells (data not shown) were efficiently lysed in the presence of this drug. Analysis by flow cytometry using anti-HLA-DR monoclonal antibody showed no change in HLA-DR surface expression after 6 h of BFA treatment (unpublished results).

Because class II presentation after viral infection was not completely inhibited by BFA, we postulated that two separate antigen processing pathways might exist in these cells: an exogenous, endosomal pathway which would be BFA insensitive and an endogenous pathway which, like class I endogenous antigen presentation, would be completely BFA sensitive. To define a primarily exogenous pathway for influenza virus we used formalin-inactivated influenza A virus to sensitize target cells. Under these conditions, there is little or no BFA effect on class II-restricted antigen presentation, whereas chloroquine abolishes this presentation, presumably by inhibiting endosomal processing of viral antigens (Fig. 2). These results are consistent with previous findings showing a lack of BFA inhibition of presentation of exogenous antigen in B cell lines¹². To isolate the endogenous pathway for influenza viral antigen presentation, viruses must be allowed to enter the cytoplasm without passing through an acidic endosomal compartment. An efficient method for direct viral fusion to the cell plasma membrane exists for viruses such as influenza A, which normally fuse with the membranes of endosomes after acidification²⁰. When virus is absorbed to the plasma membrane at 4 °C in the presence of chloroquine (so that no endocytosis or endosomal acidification

can occur), and the pH of the medium is lowered to below 5.0, rapid viral fusion takes place and productive viral infection is initiated²¹⁻²³. When live influenza A/JAP/57 virus was absorbed and fused to K4B cells using this technique, we observed efficient, chloroquine-insensitive antigen presentation to class II-restricted CTL (Fig. 3a, b). This endogenous presentation is abrogated by BFA (Fig. 3a, b). BFA does not inhibit killing of mock infected, M1 17-33 peptide pulsed cells under these conditions (data not shown).

These data demonstrate the existence of the endogenous pathway for class II-restricted cytoplasmic antigens. With remarkable similarity to the class I-restricted pathway, efficient presentation depends on cytoplasmic localization of antigen and egress of protein from the ER. By analogy with our model for the class I-restricted pathway¹², we predict that cytoplasmic antigens, or their partially degraded peptides, are transported into the ER. Here they could either bind to MHC class I or II, or be secreted or degraded. Because presentation of endogenous antigen is chloroquine-insensitive, taking place in the absence of acidic endosomal processing, these antigens are unlikely to be processed in the endosomal system. Therefore, we favour a mechanism in which processed endogenous antigen binds to class II MHC in the ER and is transported through the normal protein secretory pathway to reach the plasma membrane. Our data are also consistent, however, with a mechanism in which processed endogenous antigen enters the endosome by a route other than the normal secretory pathway, but binds only to newly synthesized class II molecules.

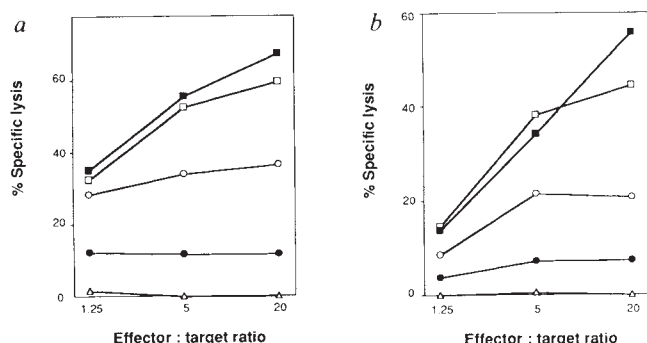


FIG. 1 BFA inhibits class II-restricted cell-mediated lysis. Lysis of K4B target cells by the HLA-DR1 restricted influenza A-specific CTL lines 130.1C6 (a) and 109.2B2 (b). K4B target cells: Uninfected (Δ); pulsed with influenza A/JAP/57 virus for 1 h ($\circ$); pulsed with M1 17-31 peptide for 1 h ($\square$); or pulsed with virus ($\bullet$) or peptide ($\blacksquare$) for 1 h and then treated with BFA. METHODS. CD4⁺ CTL lines specific for the influenza A matrix-derived peptide M1 17-31 were generated from peripheral blood lymphocytes of the HLA-DR1 positive donors 130.1C6 and 109.2B2 as described previously¹⁷. CTL activity was measured in a standard 5-h ⁵¹Cr release assay³². HLA-DR1⁺ Epstein-Barr virus transformed human B lymphoblastoid cells (K4B) were incubated overnight with ⁵¹Cr, washed, resuspended at 4×10^6 ml⁻¹ in RPMI 1640 containing 5% fetal calf serum (RPMI/5), infected with influenza A/JAP/57 (100 μ l infectious allantoic fluid per 2×10^6 cells) or incubated with M1 17-31 peptide at 50 μ g ml⁻¹ for 1 h (37 °C), washed, and resuspended at 5×10^5 ml⁻¹ in RPMI/5 $\pm$ 0.5 μ g ml⁻¹ BFA (BFA) at 37 °C and incubated for a total of 5.5 h after infection. Washed cells were plated at 5×10^3 per well in 100 μ l RPMI/5 $\pm$ 1 μ g ml⁻¹ BFA. CTL cells at the appropriate effector target ratios were added in 100 μ l RPMI/5. After 5 h the supernatants were assayed for ⁵¹Cr release and the per cent specific lysis was determined from the expression (release by CTL - spontaneous release)/(release by 1% Triton X-100 - spontaneous release) $\times$ 100.

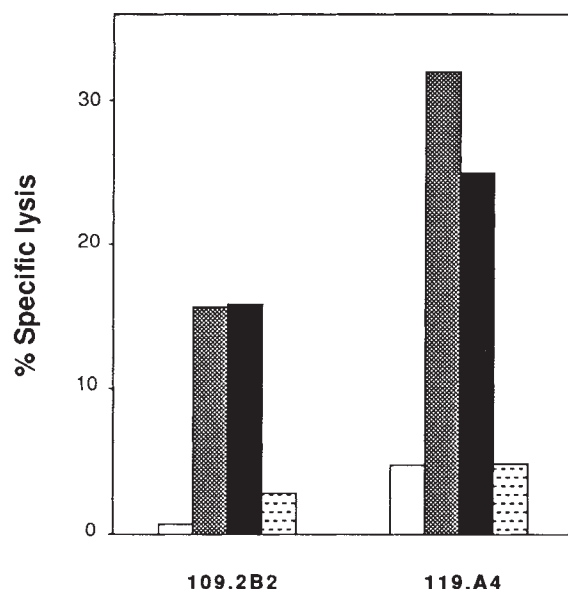


FIG. 2 Chloroquine, but not BFA, inhibits presentation of inactivated influenza virus. Lysis by the HLA-DR1 restricted CTL lines 109.2B2 and 119.A4 of K4B cells exposed to formalin-inactivated influenza A/Sichuan/X-97 virus (grey bars), exposed to inactivated virus and treated with either BFA (black bars) or chloroquine (stippled bars), or not exposed (white bars). METHODS. K4B cells were exposed to and incubated with inactivated influenza virus as with infection with live virus (see legend to Fig. 1). Instead of using 100 μ l infectious allantoic fluid, 100 μ l of a 1:100 dilution of formalin-inactivated virus stock (Connaught Labs) was used (final concentration 700 μ g whole virus per ml RPMI/5). For K4B cells treated with chloroquine, [⁵¹Cr] labelled cells were washed and resuspended in RPMI/5 containing 60 μ M chloroquine (Sigma) for 20 min at 25 °C. Diluted inactivated virus (100 μ l) was added, the cells incubated for 1 h at 37 °C, washed in RPMI/5 containing 50 μ M chloroquine, resuspended in RPMI/5 containing 5 μ M chloroquine and incubated at 37 °C for a total of 5.5 h after infection. Washed cells (RPMI/5 containing 5 μ M chloroquine) were plated in 10 μ M chloroquine and an equal volume of CTL cells (effector:target ratio was 20 for 109.2B2 and 5 for 119.A4) in RPMI/5 was added to each well, giving a final concentration of 5 μ M chloroquine during the 5 h CTL assay.

This model of class II-restricted presentation of endogenous antigen can account both for data demonstrating efficient endogenous antigen presentation of non-membrane measles virus antigens^{15,24,25}, and for the lack of endogenous presentation of influenza haemagglutinin antigens in identical B lymphoblastoid cells^{1,14}. Measles virus fuses with the plasma membrane releasing nucleocapsids and viral RNA directly into the cytoplasm²⁶, whereas influenza virus fuses with the cell membrane only after acidification of endosomes into which it has previously been delivered by endocytosis. Because this latter process is blocked by chloroquine²⁷, it has been difficult to study influenza viral antigen presentation under conditions where endosomal processing is inhibited. We have now shown that when influenza virus internal antigens and ribonucleoprotein are allowed to enter the cytoplasm of the cell directly in the same way as measles virus, efficient antigen presentation can take place only by this endogenous BFA-sensitive pathway.

These results indicate that the distinction between available antigen presentation pathways for class I and class II MHC molecules may not hold for all antigens. The growing conviction that the site of interaction of class I with its peptide is in the ER^{12,13,28} raises the question as to why class II, which obviously passes through this compartment, would not bind peptides there and thus manifest the phenotype of presenting endogenous antigen. A candidate for a molecule that might prevent the association of antigen with class II in the ER is the invariant chain that associates with newly synthesized class II in the ER and is removed only much later in a peripheral, acidic (possibly endosomal) compartment²⁹. Our results suggest, however, that DR1 can bind endogenous peptide in a manner indistinguishable from that seen for class I molecules and that the presence of invariant chain does not prevent this binding. Sekaly *et al.* have reported that the absence of invariant chain does not prevent the presentation of measles virus cytosolic antigens¹⁵. On the other hand the absence of invariant chain does seem to hamper the presentation of some exogenous antigens by the endosomal pathway^{30,31}. From the available data, we would suggest that class II MHC molecules are, like class I, able to bind endogenous

antigens in the ER pathway, and that this might take place regardless of the presence of invariant chain. It may be that the extent to which this pathway is used may vary with the particular epitope, class II molecule and cell under examination. In addition to the endogenous pathway, class II, but not class I molecules, are competent to bind antigens in the exogenous endosomal pathway. Perhaps the function of invariant chain is to provide this second option to class II. This may be achieved by its role in the intracellular sorting of class II to the endosomal compartment or by facilitating the exchange of prebound endogenous peptides for exogenous peptides. □

Received 24 August; accepted 3 November, 1989.

- Morrison, L. A., Lukacher, A. E., Braciale, V. L., Fan, D. P. & Braciale, T. J. *J. exp. Med.* **163**, 903-921 (1986).
- Germann, R. N. *Nature* **322**, 687-689 (1986).
- Sweetser, M. L., Morrison, L. A., Braciale, V. L. & Braciale, T. J. *Nature* **342**, 180-182 (1989).
- Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457-467 (1985).
- Townsend, A. R. M., Bastin, J., Gould, K. & Brownlee, G. G. *Nature* **324**, 575-577 (1986).
- Townsend, A. R. M. *et al. J. exp. Med.* **168**, 1211-1224 (1988).
- Yewdell, J. W., Bennick, J. R. & Hosaka, Y. *Science* **229**, 637-640 (1988).
- Moore, M. W., Carbone, F. R. & Bevan, M. J. *Cell* **54**, 777-785 (1988).
- Takatsuki, A. & Tamura, G. *Agric. Biol. Chem.* **49**, 899-902 (1985).
- Misumi, Y. *et al. J. Biol. Chem.* **261**, 11398-11403 (1986).
- Lippincott-Schwartz, J., Yuan, L. C., Bonifacio, J. S. & Klausner, R. D. *Cell* **56**, 801-813 (1989).
- Nuchtern, J. G., Bonifacio, J. S., Biddison, W. E. & Klausner, R. D. *Nature* **339**, 223-226 (1989).
- Yewdell, J. W. & Bennick, J. R. *Science* **244**, 1072-1075 (1989).
- Fleischer, B., Becht, H. & Rott, R. *J. Immunol.* **135**, 2800-2804 (1985).
- Sekaly, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 1209-1212 (1988).
- Rothbard, J. *et al. Cell* **52**, 515-523 (1988).
- Shimono, N., Maloy, W. L., Anderson, R. W., Biddison, W. E. & Coligan, J. E. *J. Immunol.* **143**, 2939-2947 (1989).
- Gotch, F., Rothbard, J., Howland, K., Townsend, A. & McMichael, A. *Nature* **326**, 881-882 (1987).
- Hogan, K. T. *et al. J. exp. Med.* **168**, 725-736 (1988).
- Helenius, A., Kartenbeck, J., Simons, K. & Fries, E. *J. Cell Biol.* **84**, 404-420 (1980).
- White, J., Kartenbeck, J. & Helenius, A. *J. Cell Biol.* **87**, 264-272 (1980).
- White, J., Matlin, K. & Helenius, A. *J. Cell Biol.* **89**, 674-679 (1981).
- Matlin, K., Reggio, H., Helenius, A. & Simons, K. *J. Cell Biol.* **91**, 601-613 (1981).
- Jacobson, S., Sekaly, R. P., Jacobson, C. L., McFarland, H. F. & Long, E. O. *J. Virol.* **63**, 1756-1762 (1989).
- Long, E. O. & Jacobson, S. *Immun. Today* **10**, 45-48 (1989).
- Frazer, K. B. & Martin, S. J. in *Measles Virus and Its Biology* p. 74 (Academic, London, 1978).
- Helenius, A., Marsh, M. & White, J. *J. gen. Virol.* **58**, 47-61 (1982).
- Townsend, A. R. M. *et al. Nature* **340**, 443-448 (1989).
- Blum, J. S. & Cresswell, P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3975-3979 (1988).
- Malissen, B. *et al. Cell* **36**, 319-327 (1984).
- Stockinger, B. *et al. Cell* **56**, 683-689 (1989).
- Cowan, E. P., Jelachich, M. L., Coligan, J. E. & Biddison, W. E. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5014-5018 (1987).

ACKNOWLEDGEMENTS. We thank Drs Eric Long, Juan S. Bonifacio, and Jennifer Lippincott-Schwartz for helpful discussions, Richard Turner for technical assistance and Ekaterini Perry for manuscript preparation. Inactivated influenza virus was a gift from Connaught Laboratories, Inc. and BFA was a gift from the Sandoz Company.

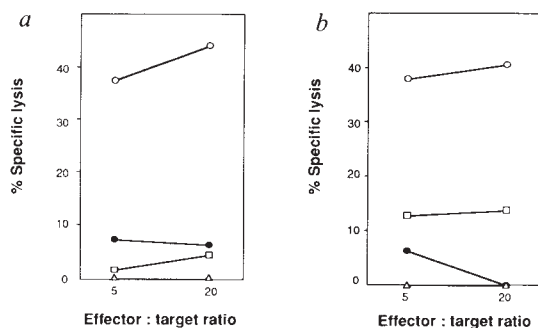


FIG. 3 BFA, but not chloroquine, blocks class II-restricted lysis after direct fusion of virus with target cell membranes. Lysis was by CTL lines 130.1C6 (a) and 109.2B2 (b). Chloroquine-treated K4B cells were uninfected (Δ); incubated with infectious influenza virus and subjected to acidic pH pulse (○); subjected to pH pulse and subsequently treated with BFA (●); or not subjected to pH pulse (□).

METHODS. K4B cells were labelled overnight with ⁵¹Cr and washed in virus-binding medium²¹ (VBM) consisting of RPMI 1640 containing 10 mM HEPES, pH 6.8, 0.2% BSA (fraction V; Sigma) and 50 μM chloroquine. The cells were resuspended in 0.5 ml VBM containing 60 μM chloroquine at 4 °C and combined with influenza A/JAP/57 virus (100 μl infectious allantoic fluid) for 1 h at 4 °C. One ml of low-pH medium (VBM containing 110 mM sodium acetate buffer, pH 4.3, and 50 μM chloroquine) or 1 ml VBM containing 50 μM chloroquine (for non-pH pulsed control) was added to cells (final pH 4.9), which were warmed to 37 °C for 90 s before adding 10 ml VBM (50 μM chloroquine) (restoring the pH to 6.6). Target cells were washed in RPMI/5 (50 μM chloroquine) ± 0.5 μg ml⁻¹ BFA and subsequently treated in the same way as the chloroquine-treated K4B cells in Fig. 2, except that 0.5 μg ml⁻¹ BFA was maintained in the BFA-treated group throughout the end of the CTL assay.

Haemopoietic colony stimulating factors promote cell survival by suppressing apoptosis

Gwyn T. Williams*, Christopher A. Smith*, Elaine Spooner†, T. Michael Dexter† & Dale R. Taylor*

* Department of Anatomy, University of Birmingham Medical School, Vincent Drive, Birmingham, B15 2TJ, UK

† Department of Experimental Haematology, Paterson Institute for Cancer Research, Christie Hospital and Holt Radium Institute, Manchester, M20 9BX, UK

THE survival, differentiation, proliferation and development of haemopoietic precursor cells and the functional activity of mature blood cells are all influenced by colony stimulating factors (CSFs)¹. As haemopoietic cells rapidly die in the absence of appropriate CSF^{2,3}, the promotion of cell survival mediated by CSFs, or growth factors, is fundamental to all the other effects exerted by these factors. This enhancement of cell survival is distinct from the stimulation of proliferation^{1,4}. Here we show that the death of haemopoietic precursor cells on withdrawal of the relevant CSF

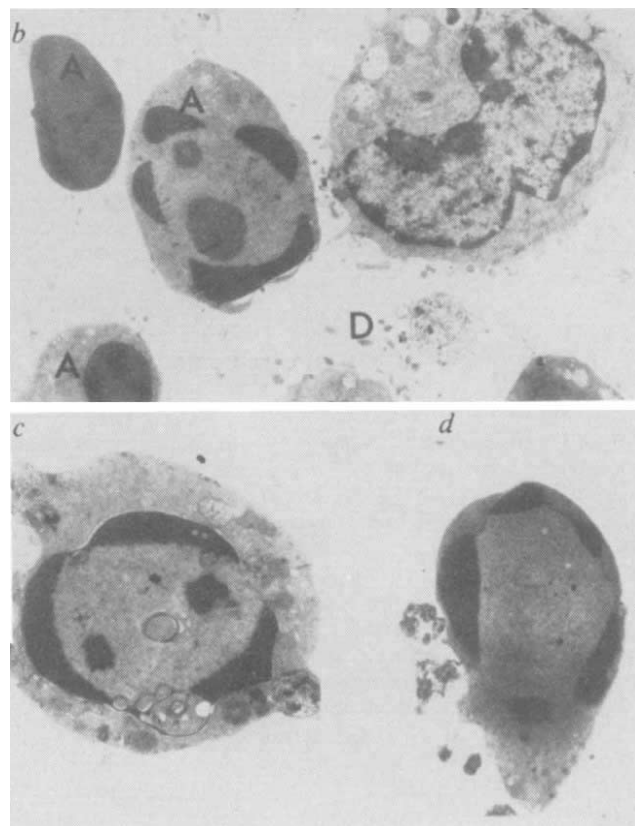
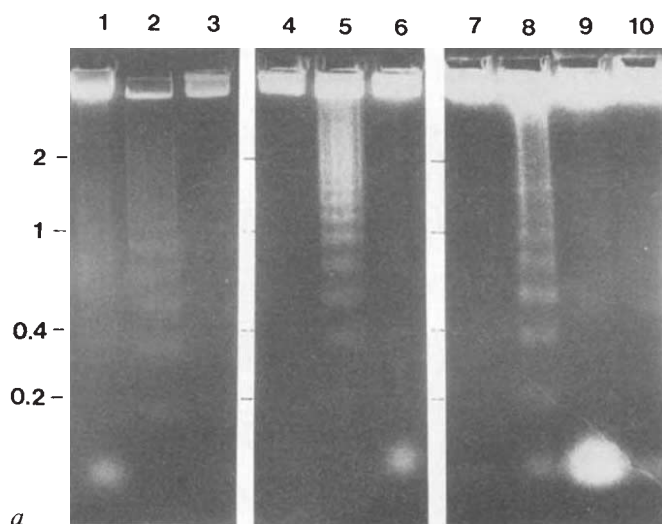


FIG. 1 Apoptosis produced by IL-3 deprivation of dependent cell lines. *a*, Total DNA from IL-3 dependent cells incubated with IL-3 shows minimal degradation after 19 h on analysis by agarose gel electrophoresis. DNA from cells incubated without IL-3 shows a clear ladder of degraded DNA bands, the sizes of which are multiples of ~ 200 bp. This pattern of DNA degradation is characteristic of apoptosis^{8,9,18,19}. The fragmentation of DNA was suppressed in those cultures to which IL-3 was added at the start of the incubation. Lanes 1–3, FDCP-Mix A4 cells: lane 1, 0 h; lane 2, 19 h without IL-3; lane 3, 19 h with IL-3. Lanes 4–6, FDCP-Mix 1 cells: lane 4, 0 h; lane 5, 19 h without IL-3; lane 6, 19 h with IL-3. Lanes 7–10, FDCP-1 cells: lane 7, 0 h; lane 8, 19 h without IL-3; lane 9, 19 h with IL-3; lane 10, 19 h with recombinant IL-3. Molecular sizes in kilobases (kb) are indicated. Material at the bottom of the gel is the product of RNase digestion. *b*, Electron micrograph of FDCP-Mix A4 cells after 19 h incubation without IL-3. Cells at different stages of apoptosis are labelled A, and cell debris labelled D. A cell showing relatively normal morphology at this time is also shown. *c*, Electron micrograph of FDCP-Mix 1 cell after 19 h without IL-3. *d*, Electron micrograph of FDCP-1 cell after 19 h without IL-3. The highly condensed chromatin seen in *b–d* is typical of apoptosis. METHODS. Cells were prepared by culture in Fischer's medium with 20% horse serum (FHS) supplemented with 10% (v/v) WEHI-3 cell conditioned

medium to provide IL-3 (ref. 13). All cells were washed with Fischer's medium alone to remove IL-3, and were resuspended in FHS without IL-3 and placed into replicate 10 ml cultures in tissue culture flasks at a final concentration of 2×10^5 cells ml^{-1} . Control cultures were supplemented with WEHI-3 cell conditioned medium as a source of IL-3 (ref. 13) or with purified rIL-3 (150 U ml^{-1} , Biogen) where stated. After 19 h incubation at 37 °C, 10^6 cells were collected for each culture and total DNA analysed by gel electrophoresis, as previously described⁸.

is due to active cell death^{5,6}, or apoptosis, indicating that CSFs promote cell survival by suppression of the process of apoptosis. The existence of a positive control mechanism regulating precursor cell survival has important implications both for the regulation of normal haemopoiesis and for tumorigenesis.

Much cell death that is physiologically significant, as, for example, in embryo morphogenesis⁷, occurs by the active process of self-destruction termed apoptosis^{5,6}. It has recently been shown that apoptosis occurs in several parts of the immune system, and in particular in the death of immature thymocytes induced through their antigen receptor complex^{8,9}, in mature interleukin-2 dependent T-cell lines on withdrawal of the growth factor¹⁰, in target cells attacked by cytotoxic T cells^{11,12} and in germinal centre B cells incubated without stimulatory monoclonal antibodies²⁸. As some of these examples involve apoptosis resulting from the absence of specific signals, we analysed the death of haemopoietic precursor cells on withdrawal of the appropriate CSF.

Haemopoietic precursor cell lines FDCP-1 (ref. 13) and FDCP-Mix (ref. 14) were derived from long-term mouse bone marrow cultures. The FDCP-1 cell line represents a granulocyte progenitor population which is blocked in the differentiation pathway. Two FDCP-Mix clones, A4 (ref. 15) and 1 (ref. 14) share many characteristics with normal haemopoietic stem cells, in particular, the ability to differentiate in response to normal regulatory signals¹⁴. The growth of these and similar cells is dependent on the addition of CSFs such as interleukin-3 (IL-3, also referred to as multi-CSF), granulocyte/macrophage colony

stimulating factor (GM-CSF) or granulocyte colony stimulating factor (G-CSF). Such 'factor-dependent' cells die rapidly when deprived of the appropriate growth factor^{16,17}. All three cell lines were maintained in liquid culture supplemented with IL-3 and horse serum, and under these conditions exhibited >95% viability and >98% primitive blast cell morphology. FDCP-1, FDCP-Mix A4 and FDCP-Mix 1 grown in medium containing IL-3 were washed and incubated in separate cultures with and without IL-3. Cultures containing IL-3 remained healthy, as expected, and extensive cell death was observed in the cultures deprived of IL-3 over the 2-day incubation period (data not shown), as reported previously^{14,16,17}. When the DNA of the incubated cells was analysed by gel electrophoresis⁸, cells from cultures deprived of the interleukin for 19 h showed the extensive degradation of DNA to oligonucleosomal fragments which is characteristic of apoptosis^{8,9,18,19} (Fig. 1*a*). This ladder of DNA fragments, which results from preferential digestion of exposed internucleosomal DNA by a nuclear endonuclease, was detectable within 6 h of withdrawal of IL-3 from FDCP-Mix A4 cells (data not shown), indicating that this was an early stage in cell death. The morphological features of apoptosis¹⁹ such as the distinctive condensation of chromatin and loss of plasma membrane microvilli, were clearly visible on electron microscopy of IL-3 deprived cells (Fig. 1*b–d*).

One of the most striking features of apoptosis which illustrates the active nature of this form of cell death is that, in many systems, it is dependent on protein synthesis in the dying cell^{7,9,10,19}. We therefore monitored cell death after IL-3 with-

drawal in the presence of the protein synthesis inhibitor cycloheximide (Fig. 2). The inhibition of the rate at which the cells died was highly significant statistically (see Fig. 2 legend) and was consistently observed in additional experiments. This clear effect of cycloheximide, which is similar to that noted in other systems, provided further evidence for apoptosis.

Some FDCP-Mix cell lines can be maintained in the presence of CSFs other than IL-3 (ref. 17 and unpublished observations). We therefore examined the effect of CSF withdrawal on FDCP-Mix cell lines C2GM (GM-CSF dependent) and 1G (G-CSF dependent). In both cases the resulting cell death was accompanied by the specific DNA degradation pattern (Fig. 3a) and changes in cell morphology (Fig. 3b, c) characteristic of apoptosis. As found for IL-3, the control cell cultures to which the CSFs were added remained healthy both with CSF-containing cell supernatants and recombinant CSFs.

These results indicate that suppression of apoptosis of the haemopoietic precursor cell lines is a general property shared by IL-3, GM-CSF and G-CSF and as such may be important in regulating the sizes of normal haemopoietic precursor populations in the bone marrow. Erythroid precursor cells responsive to the hormone erythropoietin (ERC cells), for example, continue to be produced in the bone marrow even when erythropoietin is essentially absent²⁰. Because the ERC population does not increase in size in these circumstances, it is likely that the excess ERC cells produced simply die (potentially by a mechanism involving apoptosis). Unfortunately, it has not yet been possible to abrogate production of CSFs *in vivo* using protocols similar to those adopted for the study of erythropoietin and its effects. Elevation of serum CSF levels, however, by treatment of animals and humans with these factors for several days leads to dramatic increases in the numbers of precursor cells or mature cells, or both, in the bone marrow^{21,22}. Although some of the effects are probably due to stimulation of the

proliferation of cells resident in the bone marrow, our results show that at least part of the response may be due to increased survival of the steady-state precursor population.

Vaux and co-workers⁴ have recently shown that the introduction and expression of the proto-oncogene *bcl-2* suppresses the death of IL-3 deprived FDCP-1 cells but does not produce proliferation. Together with our observation that IL-3 deprivation of FDCP-1 cells produces death by apoptosis, this

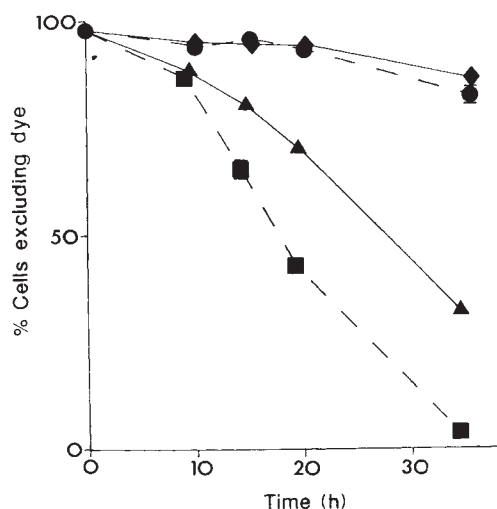


FIG. 2 Inhibition by cycloheximide of cell death following IL-3 deprivation. Of FDCP-Mix A4 cells incubated in the absence of IL-3 (■) 96% were dye-permeable after 35 h; the rate at which the cells died was inhibited by 1 μ M cycloheximide (▲). Control cultures with IL-3 (●), and with IL-3 and 1 μ M cycloheximide (◆), and standard errors of four replicate cultures wherever not encompassed by the symbols, are shown. Inhibition of cell death in cultures without IL-3 by cycloheximide was significant at the 99% level for all time points after 14 h culture.

METHODS. FDCP-Mix A4 cells were prepared and washed as in Fig. 1, and placed in 1 ml cultures in 24-well plates at a final concentration of 3.5×10^5 cells ml^{-1} . Independent cultures were used for each replicate. Appropriate cultures were supplemented with WEHI-3-cell conditioned medium containing IL-3 or 1 μ M cycloheximide, or both, as detailed above. The percentages of cells excluding nigrosin at each time point were determined using a haemocytometer.

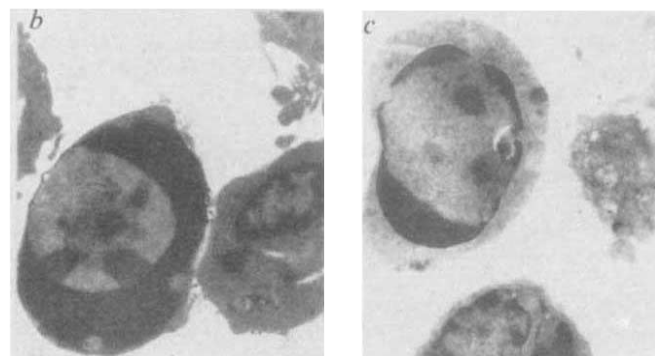
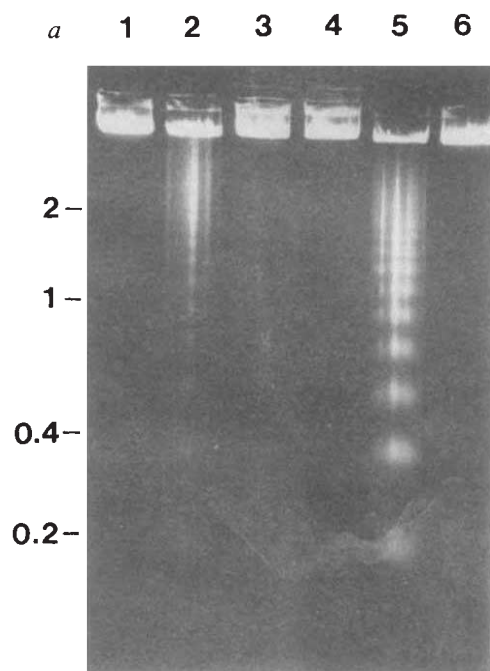


FIG. 3 Apoptosis produced by G-CSF and GM-CSF deprivation of dependent cell lines. a, Total DNA from FDCP-Mix C2GM cells after 18 h in the presence of GM-CSF, and from FDCP-Mix 1G cells after 18 h in the presence of G-CSF, shows minimal DNA degradation, but that from cells incubated without CSFs shows the clear ladder of oligonucleosome-sized DNA fragments characteristic of apoptosis^{8,9,18,19}. Lanes 1–3, FDCP-Mix C2GM cells: lane 1, 0 h; lane 2, 18 h without GM-CSF; lane 3, 18 h with GM-CSF. Lanes 4–6, FDCP-Mix 1G cells: lane 4, 0 h; lane 5, 18 h without G-CSF; lane 6, 18 h with G-CSF. b, Electron micrograph of FDCP-Mix 1G cell after 18 h without G-CSF. c, Electron micrograph of FDCP-Mix C2GM cell after 19 h without GM-CSF.

METHODS. FDCP-Mix C2GM cells were prepared by culture in FHS supplemented with 10% (v/v) lung conditioned medium as a source of GM-CSF²⁶. FDCP-Mix 1G cells were prepared by culture in Iscove's medium with 10% horse serum (IHS) supplemented with 20% (v/v) 5637 cell conditioned medium as a source of G-CSF²⁷. Cells were washed with Fischer's medium and resuspended in FHS (C2GM) or IHS (1G) and placed in replicate tissue culture flasks at a final concentration of 2×10^5 cells ml^{-1} , 10 ml per culture. Control cultures were supplemented with 50 U/ ml^{-1} (Biogen) murine recombinant GM-CSF (C2GM) or 1,000 U/ ml^{-1} (Amgen) human recombinant G-CSF (1G). Total DNA was analysed as in Fig. 1.

indicates that the *bcl-2* gene product can act, directly or indirectly, to suppress apoptosis. It is an attractive working hypothesis that such suppression of the apoptosis occurring in germinal centre B cells²⁸ could be important in the development of follicular lymphoma, with which *bcl-2* translocations are strongly associated²³⁻²⁵. □

Note added in proof: Independent studies on the IL-3-dependent cell line FDCP-2 also indicate that growth factor withdrawal produces apoptosis (T. Crompton, personal communication).

Received 27 August; accepted 2 November 1989.

1. Metcalf, D. *Nature* **335**, 27-30 (1989).
2. Paran, M. & Sachs, L. *J. Cell. Physiol.* **72**, 247-249 (1969).
3. Metcalf, D. *J. cell. Physiol.* **76**, 89-100 (1970).
4. Vaux, D. L., Cory, S. & Adams, J. M. *Nature* **335**, 440-442 (1988).
5. Kerr, J. F. R., Wyllie, A. H. & Currie, A. R. *Br. J. Cancer* **26**, 239-257 (1972).
6. Wyllie, A. H., Kerr, J. F. R. & Currie, A. R. *Int. Rev. Cytol.* **68**, 251-305 (1981).
7. Pratt, R. B. & Greene, R. M. *Devel. Biol.* **54**, 135-145 (1976).
8. Smith, C. A., Williams, G. T., Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **337**, 181-184 (1989).
9. Shi, Y., Sahai, B. M. & Green, D. R. *Nature* **339**, 625-626 (1989).
10. Duke, R. C. & Cohen, J. J. *Lymphokine Res.* **5**, 289-299 (1986).
11. Russell, J. H., Maszkowski, V. R. & Dobos, C. B. *J. Immunol.* **124**, 1100-1105 (1980).
12. Ucker, D. S. *Nature* **327**, 62-64 (1987).
13. Dexter, T. M., Garland, J., Scott, D., Scolnick, E. & Metcalf, D. *J. exp. Med.* **152**, 1036-1047 (1980).
14. Spooner, E., Heyworth, C. M., Dunn, A. & Dexter, T. M. *Differentiation* **31**, 111-118 (1986).
15. Spooner, E. *et al. Curr. Top. micro. Immunol.* **149**, 109-116 (1989).
16. Whetton, A. D. & Dexter, T. M. *Nature* **303**, 629-631 (1983).
17. Roberts, R. *et al. Nature* **332**, 376-378 (1988).
18. Wyllie, A. H. *Nature* **284**, 555-557 (1980).
19. Wyllie, A. H., Morris, R. G., Smith, A. L. & Dunlop, D. *J. Pathol.* **142**, 67-77 (1984).
20. von Wangenheim, H. R., Schofield, R., Kyffin, S. & Stein, B. *Biomedicine* **27**, 337-340 (1977).
21. Bronchud, M. H. *et al. Br. J. Cancer* **58**, 64-69 (1988).
22. Gabrilove, J. L., Jakubowski, A., Scher, H. *et al. New Engl. J. Med.* **318**, 1414-1422 (1988).
23. Tsujimoto, Y., Finger, L. R., Yunis, J. J., Nowell, P. C. & Croce, C. M. *Science* **226**, 1097-1099 (1984).
24. Bakshi, A. *et al. Cell* **41**, 899-906 (1985).
25. Cleary, M. L., Smith, S. D. & Sklar, J. *Cell* **47**, 19-28 (1986).
26. Burgess, A. W., Camakaris, J. & Metcalf, D. *J. bio. Chem.* **252**, 1998-2003 (1977).
27. Welte, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 1526-1530 (1985).
28. Liu, Y. J. *et al. Nature* (in the press).

ACKNOWLEDGEMENTS. We thank A. Murdoch for technical assistance, A. Orchard for photography and F. Farzaneh, J. J. T. Owen, E. J. Jenkinson, G. Brown and J. Lord for their comments on the manuscript. This work was funded by the UK Medical Research Council, the Wellcome Trust and the Research Corporation Trust (G.T.W.), and the UK Cancer Research Campaign (E.S. and T.M.D.).

G_o protein as signal transducer in the pertussis toxin-sensitive phosphatidylinositol pathway

Thomas M. Moriarty, Elena Padrell, Donna J. Carty, Gila Omri, Emmanuel M. Landau & Ravi Iyengar*

Departments of Psychiatry and Pharmacology, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, New York 10029, USA and Department of Psychiatry, Bronx Veterans Administration Medical Center, 130 West Kingsbridge Road, Bronx, New York 10468, USA

RECEPTORS stimulating phospholipase C do so through heterotrimeric GTP-binding proteins^{1,2} to produce two second messengers, inositol 1,4,5-trisphosphate (InsP₃) and diacylglycerol¹. In spite of the detailed understanding of phospholipase C structure³ and phosphatidyl inositol signalling¹, the identity of the GTP-binding protein involved is so far unknown. To address this issue, we have used the *Xenopus* oocyte in which muscarinic receptors couple to phospholipase C through a pertussis toxin-sensitive GTP-binding protein². In this cell, InsP₃ mobilizes intracellular Ca²⁺ to evoke a Cl⁻ current. The magnitude of this Cl⁻ current is proportional to the amount of InsP₃ in the cell⁴, and therefore can be used as an assay for InsP₃ production. We report here that the activated α -subunit of the GTP-binding protein G_o, when directly injected into oocytes, evokes a Cl⁻ current by mobilizing Ca²⁺ from intracellular InsP₃-sensitive stores. We also show that holo-G_o, when injected into oocytes, can specifically enhance the

muscarinic receptor-stimulated Cl⁻ current. These data indicate that G_o can serve as the signal transducer of the receptor-regulated phospholipase C in *Xenopus* oocytes.

To determine whether any of the known pertussis toxin-sensitive α -subunits evokes the InsP₃-dependent Cl⁻ current, the four main pertussis toxin-sensitive GTP-binding proteins (G proteins) were purified. G_o and G_{i-1} were purified from bovine brain. G_{i-2} and G_{i-3} were prepared from human erythrocytes. The identities of the purified G proteins were verified by quantitative immunoblotting (Fig. 1a). These G proteins were activated with GTP- γ -S and Mg²⁺. The α -subunits activated by GTP- γ -S were purified free of $\beta\gamma$ -subunits and unliganded GTP- γ -S by ion-exchange chromatography. When injected into *Xenopus* oocytes, the activated G_o subunit (G_o^{*}) evoked a rapid response, whereas the three activated G_i subunits had no effect. Injection of unliganded GTP- γ -S in concentrations equivalent to those of the activated α -subunits did not stimulate a response (Fig. 1b). A silver-staining profile of the G_o^{*} preparation that evoked a response is shown (Fig. 1c).

The depolarizing response was not evoked if the G_o^{*} was boiled (2 min at 100 °C) before injection (Fig. 2, a compared with b). The time course and waveform of the G_o^{*}-evoked response (Fig. 2a) was very similar to the muscarinic receptor-evoked response (compare Fig. 2a and Fig. 4 inset), which is known to be an InsP₃-mediated Ca²⁺-dependent Cl⁻ current^{5,6}. Like the receptor-evoked response, the G_o^{*}-evoked depolarization was carried primarily by Cl⁻ ions⁷ (Fig. 2c). Injection of EGTA into the cell abolished the G_o^{*}-evoked response indicating that this response was dependent on mobilized intracellular Ca²⁺ as is the receptor-evoked response⁶ (Fig. 2d). We then determined whether the Ca²⁺ was mobilized from InsP₃-sensitive stores by studying the effect of preinjection of InsP₃ on the G_o^{*}-evoked response. This experiment is based on previous observations that injection of InsP₃ into oocytes renders the system unresponsive to subsequent stimulation by receptor activation or InsP₃ injection⁸. After a large depolarizing response to injection of InsP₃, the oocyte was unresponsive to a subsequent injection of InsP₃ or G_o^{*} (Fig. 3a, b). This lack of response is not due to inactivation of the Ca²⁺-sensitive Cl⁻ channels, because multiple injections of Ca²⁺ evoke the Cl⁻ current repeatedly (data not shown). These experiments show that G_o^{*} uses the InsP₃-sensitive Ca²⁺ stores to evoke the Cl⁻ current, probably by stimulating InsP₃ production.

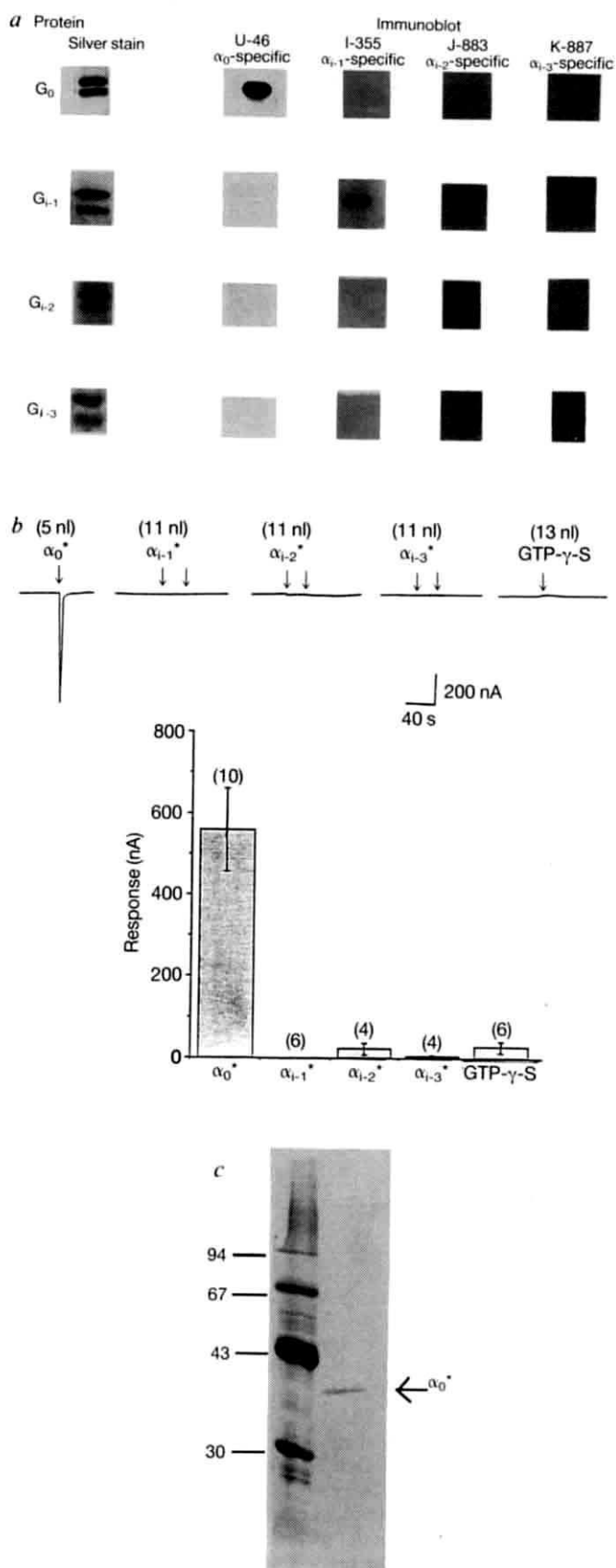
These results indicate that G_o may be able to couple muscarinic receptors to phosphatidylinositol turnover (Figs 1-3). Therefore we tested which holo-G protein could increase muscarinic stimulation of the InsP₃-mediated Cl⁻ current. The assay used is based on observations in other G protein-coupled systems, such as the G_s regulation of adenylyl cyclase, in which addition of exogenous G proteins to membranes that contain a normal complement of G protein results in enhanced stimulated activity (ref. 9, and R. T. Premont and R.I., unpublished observations). Purified holo-G-proteins were injected into groups of oocytes. After 30 min, individual cells were voltage-clamped and tested for their responsiveness to acetylcholine. Figure 4a shows representative current traces from a control cell (left) and a cell injected with G_o (right) from the same experiment. Figure 4b represents a composite of three experiments. Injection of G_o resulted in a substantial increase in the muscarinic response, whereas injection of the three G_i proteins resulted in no significant increase compared with non-injected controls. The waveform of the evoked Cl⁻ current was not altered; only the amplitude of the current was increased (Fig. 4a).

The data presented here show that the GTP-binding protein G_o is specifically capable of coupling muscarinic receptors to the production of InsP₃, and that the activated α -subunit of G_o can directly stimulate the InsP₃ pathway. The three G_i proteins or their α -subunits did not have any measurable stimulatory effects on the Cl⁻ current. The most probable site of G_o action is between the receptor and the effector. The data presented

* To whom correspondence should be addressed.

FIG. 1 Direct injection of the GTP- γ -S-activated α_0 -subunit of G_0 evokes a depolarizing current in *Xenopus* oocytes, whereas injection of activated α -subunits of G_{i-1} , G_{i-2} or G_{i-3} has no effect. **a**, Silver-staining and immunoblotting profile of the G proteins used. **b**, Activated α -subunits of the four main pertussis toxin-sensitive G proteins were pressure-injected directly into single voltage-clamped oocytes. Volumes of solutions injected are indicated. Protein concentrations of the solutions injected (nM): α_0 , 58; α_{i-1} , 60; α_{i-2} and α_{i-3} , 50. Bar graph summarizes the results from several cells. The current traces above the bar graph show representative responses to each of the activated α -subunits tested. GTP- γ -S (50 nM, 13 nl), same buffer as the α -subunits, was also injected into a separate group of oocytes; these experiments were done to rule out the possibility that the positive result seen with G_{α_0} was due to unliganded GTP- γ -S in the injector pipette. **c**, Silver-staining profile of the GTP- γ -S activated $G_{\alpha_0}^*$ after purification on hydroxylapatite column and concentration on a Centricon concentrator. Positions of relative molecular mass (M_r) standards are shown ($M_r \times 10^{-3}$).

METHODS. G protein purification. G_0 and G_{i-1} were purified from bovine brain and G_{i-2} and G_{i-3} were purified from human erythrocytes essentially as described previously¹⁹, with the following modifications. The mixed G proteins obtained after the second DEAE-Sephacel elution were diluted and loaded onto a Mono-Q column connected to a Pharmacia FPLC system and eluted with a shallow salt gradient (100–300 mM NaCl over 60 min) in a buffer containing 10 mM Na-HEPES buffer, 1 mM EDTA, 20 mM β -mercaptoethanol, 11 mM CHAPS (3-[(3-cholamidopropyl)-dimethylamino]-1-propane sulphate) and 0.1% Lubrol at pH 7.5. The peak tubes containing the resolved G proteins were identified by Coomassie-blue staining of proteins resolved by SDS-PAGE. The identities of the individual G proteins were obtained by immunoblotting with α -subunit sequence-specific antisera (gift of Drs S. Mumby and A. G. Gilman). Repeated chromatography (2–4 times) of the peak tubes over Mono-Q was required to obtain apparently homogenous individual G proteins as assessed by immunoblotting. The purified G proteins were stored in a buffer containing 25 mM Na-HEPES, 1 mM EDTA, 20 mM β -mercaptoethanol, 30% (v/v) ethyleneglycol, 0.1% Lubrol and 100 mM NaCl in a final concentration of 50–60 μ g ml⁻¹. For the determination of the protein profile, 200–500 ng of each G protein was subjected to SDS-PAGE. The gels were washed and stained with silver. Immunoblotting was performed with four sequence-specific antisera that are specific for α_0 , α_{i-1} , α_{i-2} and α_{i-3} subunits. Specificity of these antisera was verified by studying cross-reactivity with recombinant α -subunit fusion proteins expressed in *Escherichia coli*. For immunoblotting, 2–4 μ g of purified holo-G-proteins were resolved by SDS-PAGE, transferred to nitrocellulose paper and analysed with the appropriate antisera. By immunoblot analysis the proteins were found to be at least 97% pure: See ref. 20 for details of quantitative immunoblotting. Electrophysiology: Stage V and VI oocytes from mature female *X. laevis* were removed surgically under anaesthesia (0.15% tricaine immersion) and placed in ND96 frog Ringer's solution (96.0 mM NaCl, 2.0 mM KCl, 1.8 mM CaCl₂, 5.0 mM Na-HEPES, pH 7.50 $\pm$ 0.02). The follicular tissues surrounding the oocytes were enzymatically removed by gentle agitation in collagenase (2 mg ml⁻¹) containing Ca²⁺-free ND96. The oocytes were incubated in ND96 with pyruvate (2.5 mM), penicillin (100 U ml⁻¹) and streptomycin (100 μ g ml⁻¹). Oocytes were assayed 3–7 days after collection by a standard two-electrode voltage-clamp technique as previously described¹⁶. Individual oocytes were placed in a 0.5 ml bath constantly perfused with ND96 at room temperature. The oocyte was voltage-clamped at -70 mV with two KCl (3 M)-filled electrodes (0.5–2.0 M Ω). Amplifier outputs were recorded directly on chart paper and by an x-y axes plotter. Direct injection experiments were done with a picospritzer injector apparatus as previously described¹⁶. Oocytes were impaled with the injection electrode in ND96 solution. After impalement, the external bathing medium was changed to a Ca²⁺-free medium (Ca²⁺-free ND96 with 10 mM MgCl₂, 0.1 mM EGTA), in which injection of the protein and electrophysiological measurements were done. Each injection consisted of several very brief pulses. The site of injections was <100 μ m from the inner surface of the plasma membrane as measured when pulling the injector pipette out. G-protein α -subunit activation: For G_0 and G_{i-1} , 20 μ g G protein was incubated in a solution containing 25 mM Na-HEPES buffer, 1 mM EDTA, 20 mM β -mercaptoethanol, 30% (v/v) ethyleneglycol, 0.1% Lubrol and 100 mM NaCl, 10 μ M ³⁵S-labelled-GTP- γ -S ($\sim 3 \times 10^6$ c.p.m.) and 25 mM MgCl₂ (pH 8.0) for 60 min at 32 $^{\circ}$ C (final volume, 500 μ l). After incubation, the protein was loaded onto a 100- μ l hydroxylapatite column. The column was washed extensively (~ 500 ml) with 25 mM Na-HEPES buffer, 20 mM β -mercaptoethanol, 100 mM NaCl and 2 mM MgCl₂ at pH 8.0. The protein was eluted in the same buffer containing 200 mM potassium phosphate. The samples containing the protein were identified by counting, then pooled, diluted tenfold and concentrated to 200 μ l with a Centricon concentrator. The protein concentration was found to be the same by both silver-staining and ³⁵S-labelled-GTP- γ -S counting, indicating that all of the GTP- γ -S was liganded. The concentration of the protein was



calculated by counting bound ³⁵S-labelled-GTP- γ -S. Activated α_{i-2} and α_{i-3} were the gift of Drs Juan Codina and Lutz Birnbaumer and were prepared as previously described²¹. All three α_i subunits were biologically active as judged by the capability to activate the atrial K⁺ channel²².

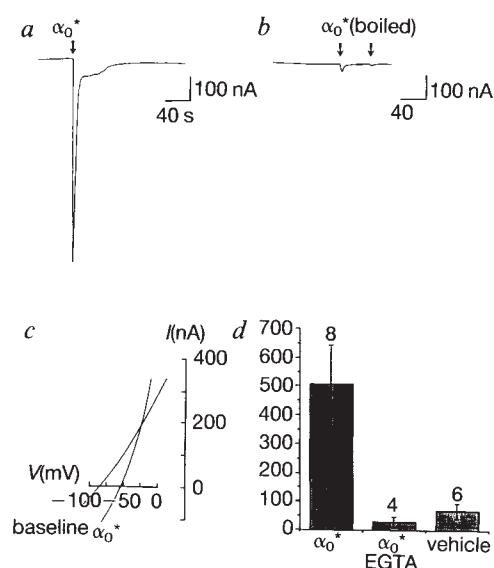


FIG. 2 Direct injection of GTP- γ -S-activated G_0 α -subunit ($G_{\alpha 0}^*$) evokes a Ca^{2+} -dependent Cl^- current. *a*, Direct pressure injection of $G_{\alpha 0}^*$ into a voltage-clamped oocyte resulted in a rapid depolarization which resembled the acetylcholine-evoked response. A representative response is shown (6 nl of 50 nM $G_{\alpha 0}^*$). *b*, Direct injection of a boiled aliquot of the same $G_{\alpha 0}^*$ as in Fig. 2*a* (9 nl and 7 nl). For experiments such as those shown in Fig. 2*a*, *b*, the experimenter did not know the identity of the injected substances. *c*, The ionic nature of the $G_{\alpha 0}^*$ -evoked conductance was examined by the ramp method⁷, in which a ramp change in voltage applied to a voltage-clamped oocyte gives the current-voltage ($I-V$) characteristics of the membrane at that moment. The $I-V$ curve is plotted directly on an $x-y$ plotter. The crossover point of an $I-V$ curve obtained at rest (baseline) and during an evoked current ($G_{\alpha 0}^*$) gives the reversal potential of the response. Voltage ramp analysis of the $G_{\alpha 0}^*$ -evoked response shows a reversal potential of -23 ± 2 mV ($n=3$). Because the equilibrium potential for Cl^- in *Xenopus* oocytes is ~ -25 mV^{6,23,24}, these results indicate that the current evoked by $G_{\alpha 0}^*$ is carried primarily by Cl^- ions. A single ramp experiment is shown. *d*, The Ca^{2+} dependence of the $G_{\alpha 0}^*$ -evoked Cl^- current was probed by pre-injecting cells with the Ca^{2+} -chelator EGTA so that the final concentration was 10^{-4} M. The bar graph compares three groups of cells: control cells in response to $G_{\alpha 0}^*$; cells pre-injected with EGTA in response to $G_{\alpha 0}^*$; and a group of cells injected with the protein storage vehicle. The $G_{\alpha 0}^*$ injections were $\sim 4-7$ nl of 50 nM $G_{\alpha 0}^*$ per oocyte. The vehicle injections were 5–10 nl storage buffer per oocyte.

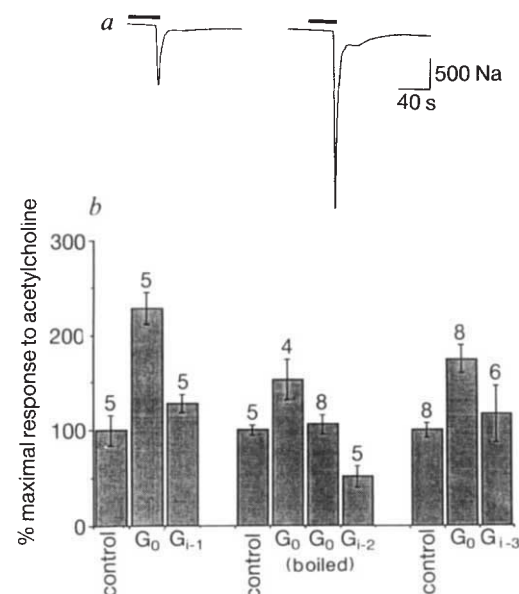


FIG. 4 G_0 , but not G_{i-1} , G_{i-2} or G_{i-3} , increases the amplitude of the muscarinic receptor-stimulated Cl^- current. *a*, Representative current traces from a voltage-clamped control cell (left) and a cell injected with holo- G_0 (right) in response to acetylcholine (1 μ M). The small bar above each trace indicates the duration of acetylcholine application. Acetylcholine was applied by superfusion. Holo-G proteins were injected with a Drummond microinjector as previously described^{2,6}. The G proteins in storage buffer were diluted sevenfold in 100 mM KCl, so as to give a 50–60 nM G-protein solution. This diluted solution (40 nl) was injected into an oocyte 30 min before the electrophysiological measurements. The amplitude of the Cl^- current was recorded at the peak of depolarization. *b*, Four pertussis toxin-sensitive G proteins were studied. Three experiments are shown comparing the acetylcholine responses of groups of oocytes that were injected with either holo- G_0 , G_{i-1} , G_{i-2} or G_{i-3} . The control cells were not injected. A group of cells injected with a heat-inactivated G_0 (middle) is included for comparison with non-injected controls. (Significant augmentation, according to t -test: control versus G_0 for the three experiments— $P < 0.005$, $P < 0.025$, $P < 0.005$; control versus G_0 (boiled)— $P < 0.4$ (not significant); control versus G_{i-1} — $P < 0.1$ (not significant); control versus G_{i-2} — $P < 0.005$ (significant reduction); control versus G_{i-3} — $P < 0.4$ (not significant).) In five experiments, the G proteins never showed any facilitatory effects. The G_0 protein was tested nine times using two different preparations of G_0 . Significant increases were observed in eight of the nine experiments. The extent of increase was variable (35–130%, mean $72 \pm 11\%$; $n=8$).

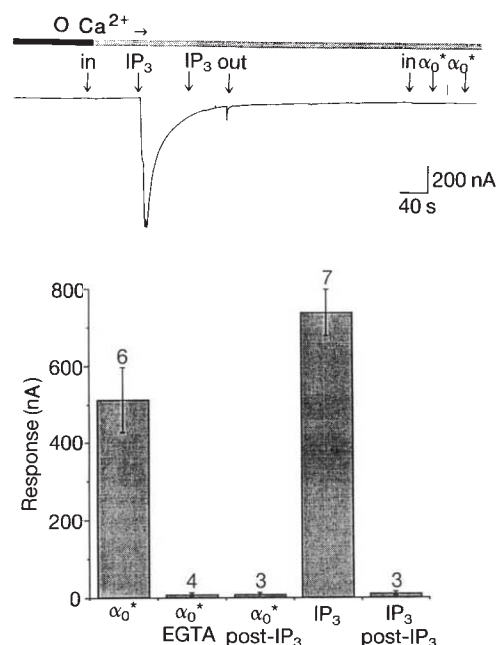


FIG. 3 Direct injection of $G_{\alpha 0}^*$ and $InsP_3$ evoke the Ca^{2+} -dependent Cl^- current through the same mechanism. A voltage-clamped oocyte was impaled with an $InsP_3$ -containing injector pipette (in). After switching the external medium from ND96 to Ca^{2+} -free ND96 (bar above trace), $InsP_3$ (2.6 pmol) was injected (IP_3) and a large depolarization resulted. Subsequent injection of $InsP_3$ (2.2 pmol) (IP_3) did not yield a response. The injector pipette was then removed (out), and a new pipette containing $G_{\alpha 0}^*$ (50 nM) was inserted (in). Injection of $G_{\alpha 0}^*$ failed to elicit a response ($G_{\alpha 0}^*$) (18 nl and 7 nl). The $G_{\alpha 0}^*$ -containing injector pipette was then impaled into another cell to confirm that the sample of $G_{\alpha 0}^*$ was active. The results for a series of experiments are summarized in the bar chart. The numbers above the bars indicate number of cells tested for each condition.

here and our previous observations on the inhibitory role of G protein $\beta\gamma$ subunits in the activation of *Xenopus* oocyte phospholipase C² indicate that G_O functions in the phospholipase C pathway in a manner similar to that of G_s in the adenylyl cyclase system, and G_i in the retinal phosphodiesterase system¹⁰.

Although G_O is generally considered to be a brain-specific G protein because it occurs there most abundantly¹¹, it has also been found in non-neuronal tissues. The complementary DNA for the G_O α -subunit has been isolated from a *Xenopus* oocyte cDNA library¹². The presence of G_O has also been demonstrated by messenger RNA hybridization and immunohistochemistry in such diverse tissues as islets of Langerhans¹³, heart and testis¹⁴ and adipocytes¹⁵, raising the possibility that it could be one of the transducers for the pertussis toxin-sensitive phospholipase

C. But not all receptor-regulated phospholipase C responses are sensitive to pertussis toxin¹⁰. The pertussis toxin-insensitive pathway is not yet well understood. It is possible that there are two distinct phospholipase Cs sensitive to G protein, or that a single G protein-regulated phospholipase C is modulated by G proteins both sensitive and insensitive to pertussis toxin. The pertussis toxin-insensitive pathway also seems to involve a heterotrimeric G protein, because it is attenuated by the injection of excess $\beta\gamma$ subunits¹⁶. Receptors that use the pertussis toxin-insensitive pathway in one cell type can use the pertussis toxin-insensitive pathway in other cells^{16,17}, indicating the possibility of a close relationship between the two G proteins¹⁸. The data described here predict that there is a close homologue of G_O that is pertussis toxin-insensitive and functions in the pertussis toxin-insensitive pathway. □

Received 8 September; accepted 5 November 1989.

- Berridge, M. J. & Irvine, R. F. *Nature* **341**, 197–204 (1989).
- Moriarty, T. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 8865–8869 (1988).
- Rhee, S. G., Suh, P.-G., Ryu, S.-H. & Lee, S. Y. *Science* **244**, 546–550 (1989).
- Gillo, B., Lass, Y., Nadler, E. & Oron, Y. *J. Physiol.* **392**, 349–361 (1987).
- Oron, Y., Dascal, N., Nadler, E. & Lupu, M. *Nature* **313**, 141–143 (1985).
- Dascal, N. *CRC Crit. Rev. Biochem.* **22**, 317–387 (1987).
- Dascal, N., Landau, E. M. & Lass, Y. *J. Physiol.* **352**, 551–574 (1984).
- Sugiyama, H., Ito, I. & Hirono, C. *Nature* **325**, 531–533 (1987).
- Sternweis, P. C. & Gilman, A. G. *J. biol. Chem.* **254**, 3333–3340 (1979).
- Gilman, A. G. *Rev. Biochem.* **56**, 615–649 (1987).
- Sternweis, P. S. & Robishaw, J. D. *J. biol. Chem.* **259**, 13806–13813 (1984).
- Olate, J. *et al. FEBS Lett.* **244**, 188–192 (1989).
- Terashima, T., Katada, T., Oinuma, M., Inoue, Y. & Ui, M. *Brain Res.* **417**, 190–194 (1987).
- Mumby, S., Pang, I., Gilman, A. G. & Sternweis, P. C. *J. biol. Chem.* **263**, 2020–2026 (1988).

- Rouot, B. *et al. Biochem. J.* **260**, 307–310 (1989).
- Moriarty, T. M. *et al. J. biol. Chem.* **264**, 13524–13530 (1989).
- Ashkenai, A., Peralta, E. G., Winslow, J. W., Ramachandran, J. & Capon, D. J. *Cell* **56**, 487–493 (1989).
- Masters, S. B. *et al. Science* **241**, 448–451 (1988).
- Codina, J. *et al. J. biol. Chem.* **259**, 5874–5886 (1984).
- Premont, R. T., Buku, A. & Iyengar, R. *J. biol. Chem.* **264**, 14960–14964 (1989).
- Codina, J., Yatani, A., Grenet, D., Brown, A. M. & Birnbaumer, L. *Science* **236**, 442–445 (1987).
- Yatani, A. *et al. Nature* **336**, 680–682 (1988).
- Kusano, K., Miledi, R. & Stinnakre, J. *J. Physiol.* **328**, 143–170 (1982).
- Barish, M. E. *J. Physiol.* **342**, 309–325 (1983).

ACKNOWLEDGEMENTS. This research was supported by NIH grants (R.I.) and Veterans Administration Merit Grant (E.M.L.). T.M.M. is a trainee of the Medical Scientist Training Program. D.J.C. is a postdoctoral trainee supported by NIDA.

Induction of a novel epidermal growth factor-secreting cell lineage by mucosal ulceration in human gastrointestinal stem cells

Nicholas A. Wright, Christine Pike & George Elia

ICRF Laboratories, Lincoln's Inn Fields, London WC2 3PN, UK

EPIDERMAL growth factor, and its human homologue urogastrone (EGF/URO)¹, are secreted by the gut-associated salivary and Brunner's glands^{2,3}. Recombinant EGF/URO is a powerful stimulator of cell proliferation and differentiation in the rodent^{4–7} and neonatal human⁸ intestine. But EGF/URO is not absorbed from the adult gut^{9,10} and has no action when given through the gut lumen⁶; thus the role of secreted EGF/URO is unknown. We now report that ulceration of the epithelium anywhere in the human gastrointestinal tract induces the development of a novel cell lineage from gastrointestinal stem cells. This lineage initially appears as a bud from the base of intestinal crypts, adjacent to the ulcer, and grows locally as a tubule, ramifying to form a new small gland, and ultimately emerges onto the mucosal surface. The lineage produces neutral mucin, shows a unique lectin-binding profile and immunophenotype, is nonproliferative, and contains and secretes abundant immunoreactive EGF/URO. We propose that all gastrointestinal stem cells can produce this cell lineage after mucosal ulceration, secreting EGF/URO to stimulate cell proliferation, regeneration and ulcer healing. This cell lineage is very commonly associated with gastrointestinal mucosal ulceration, and we conclude that a principal *in vivo* role for EGF/URO is to stimulate ulcer healing throughout the gut through induction of this cell lineage in the adjacent mucosa.

Chronic ulceration of the gastrointestinal mucosa is frequent in man, occurring particularly in Crohn's disease and peptic ulcer disease. We have studied freshly-frozen and formalin-fixed tissue from 56 patients with Crohn's disease¹¹ (multiple sites were analysed in ulcerated small intestine (45 cases) and colon

(21 cases)), and also from 66 patients with chronic gastric and duodenal peptic ulcers¹². In the mucosa adjacent to ulcers, small buds appear at the bottom of the intestinal crypts (Fig. 1a). These buds are composed of columnar cells with basal nuclei; the cytoplasm contains large amounts of neutral mucin, unlike the intestinal mucous cells (goblet cells), which contain mainly acid sialomucin. The buds develop into a new small gland, still maintaining its connection with the parent crypt (Fig. 1b). Having reached a certain size, the new gland develops a tubule which, in the small intestine, grows as a duct up the connective tissue core of an adjacent villus (Fig. 1c–f), and emerges onto the villus surface through a stoma (Fig. 1f) providing access to the intestinal lumen for the secretion of the new gland. Additionally, the cells from the duct migrate directly onto the villus surface, often covering it (Fig. 1g). On the villus, these cells maintain their singular morphology, and are quite different from adjacent enterocytes and goblet cells (Fig. 1g). These morphogenetic events are summarized diagrammatically in Fig. 2.

The phenotype of the cells was investigated by orthodox histochemistry, lectin-binding characteristics and immunocytochemistry using antibodies against several differentiation antigens. The cell lineage is phenotypically different from other small intestinal lineages. Thus only cells on the villus surface stain with antibody 3B10, which recognizes a glycoprotein of relative molecular mass 150,000 that is related to carcinoembryonic antigen¹³ (Fig. 3); cells in the gland area are negative for staining, as are enterocytes and goblet cells. Most cells in the lineage are positive for staining with SM3, an antibody raised against the mucin core protein¹⁴, whereas HMFG1 and HMFG2 antibodies, which recognize overlapping sequences in the mucin core protein¹⁵, show an interesting pattern: the cells lining the gland are HMFG1 positive and HMFG2 negative, whereas the cells on the villus surface are HMFG2 positive and HMFG1 negative, that is, they show complementary staining. Again, indigenous intestinal cell lineages are negative for staining with these monoclonal antibodies. Moreover, *Lens culinaris* (LCA), a lectin which identifies binding sites for α -mannose, α -galactose and *N*-acetylgalactosamine (in that order of preference), binds to the cell membranes of the surface cell but not cells in the gland area. These observations indicate that the lineage

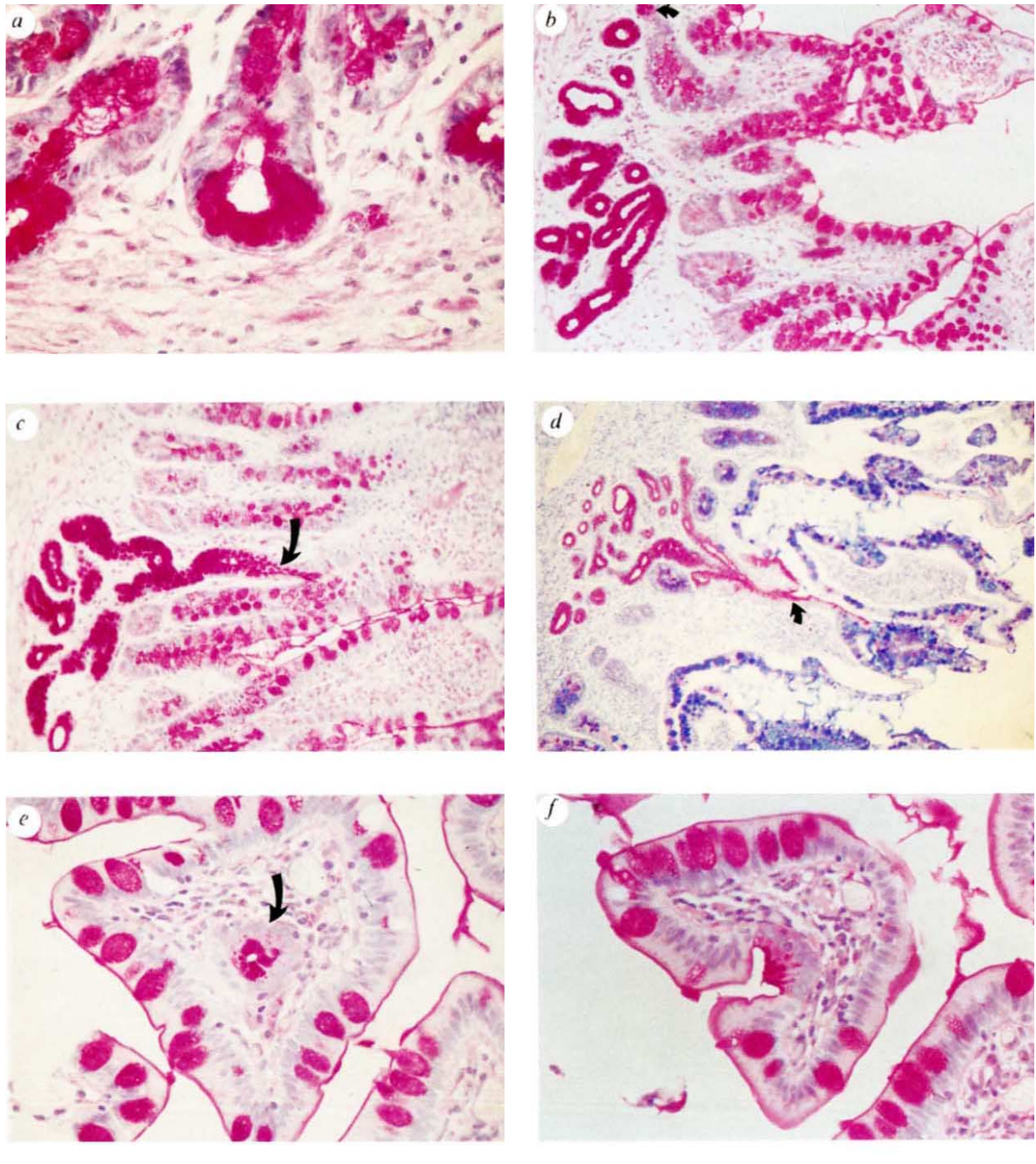


FIG. 1 *a*, The bases of intestinal crypts in the ileum from a patient with Crohn's disease, stained with the diastase periodic acid-Schiff (D/PAS) method, showing the origins of the induced cell lineage. Darkly staining buds are seen growing out from the crypt bases; the cytoplasm of the cells is intensely D/PAS-positive, reflecting the neutral mucin content. *b*, A more advanced stage of development. The tubules are now growing in the lamina propria, forming in effect a new gland between the surface intestinal villi and the underlying muscularis mucosae. Serial sections of this gland show the tubules to be connected to the parent intestinal crypt which feeds cells into the gland. *c*, Eventually the gland develops a duct (arrow) which grows up the connective tissue core of a nearby villus (*d*) to reach the luminal surface. In cross sections of villi (*e*), this duct can be seen as a tubule (arrow) occupying the centre of the villus core, which opens onto the surface (*f*). *g*, Cross-section of a villus closer to the tip, showing that the new cell lineage also migrates onto the villus surface (arrow). The distinct morphological difference between the surface enterocytes and goblet cells can be seen.

acquires new antigens as it migrates, and also changes the nature of its secretory glycoconjugates by adding new carbohydrate groups to the core protein. This suggests that the cells differentiate as they pass through the tubules onto the surface.

Both rabbit and sheep anti-human EGF/URO antibody shows abundant immunoreactive EGF/URO in cells in the gland area and in the secretion (Fig. 4); the staining appears granular, as

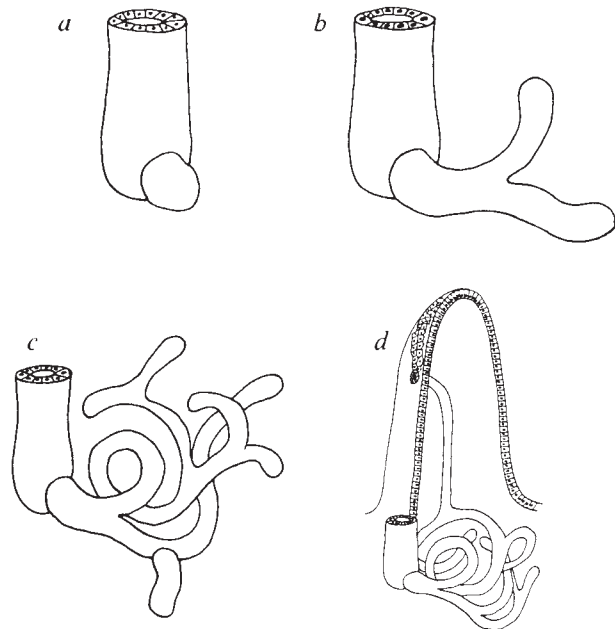


FIG. 2 A diagrammatic representation of the morphogenesis of the cell lineage: a bud appears in the base of the crypt, within the stem cell zone (a); the resulting tubules then grow in the lamina propria (b) to form a new gland (c), which eventually communicates with the surface by a duct, which pours secretion into the intestinal lumen and feeds cells onto the surface (d). For three-dimensional analysis, serial sections were prepared of the earliest stages and of glands of various sizes. These were traced onto paper using a drawing tube, the drawings stuck onto polystyrene tiles, cut out with a hot wire, and glued together. Several examples of each stage (a-d) were studied.

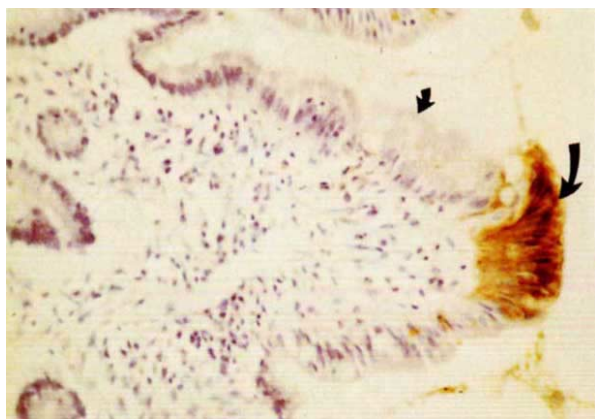


FIG. 3 The tip of an ileal villus showing a group of the cells occupying the very tip, expressing the 3B10 antigen, an M_r 150K glycoprotein related to carcinoembryonic antigen (large arrow). The surrounding villous enterocytes and goblet cells are negative (small arrow), as are the glandular portions of the cell lineage in the lamina propria. Paraffin-embedded sections were de-waxed and taken down to 100% alcohol, and endogenous peroxidase was blocked with 30% hydrogen peroxide for 15 min. Sections were incubated for 35 min with 3B10, an IgG mouse monoclonal antibody raised against normal human colon^{1,3}, washed, and incubated with biotinylated rabbit anti-mouse antibody. After 35-min incubation in avidin-biotin complex, the sections were incubated for 2 min in peroxidase substrate (diaminobenzidine, PBS, in addition to 30% hydrogen peroxide) and counterstained with haematoxylin.

seen in other EGF/URO-secreting glands³. Some acini are only faintly positive, especially in the more superficial areas, as is expected when staining for a secreted molecule, and surface cells are usually negative. The cells in basal crypt buds and small glands are faintly stained or negative for staining, indicating that the ability to secrete EGF/URO is acquired later in the lineage.

It is singular that mitotic activity in this lineage is conspicuous by its absence, even in the early basal buds (Fig. 1a), in sharp contrast to the parent crypt and adjacent crypts. Moreover the

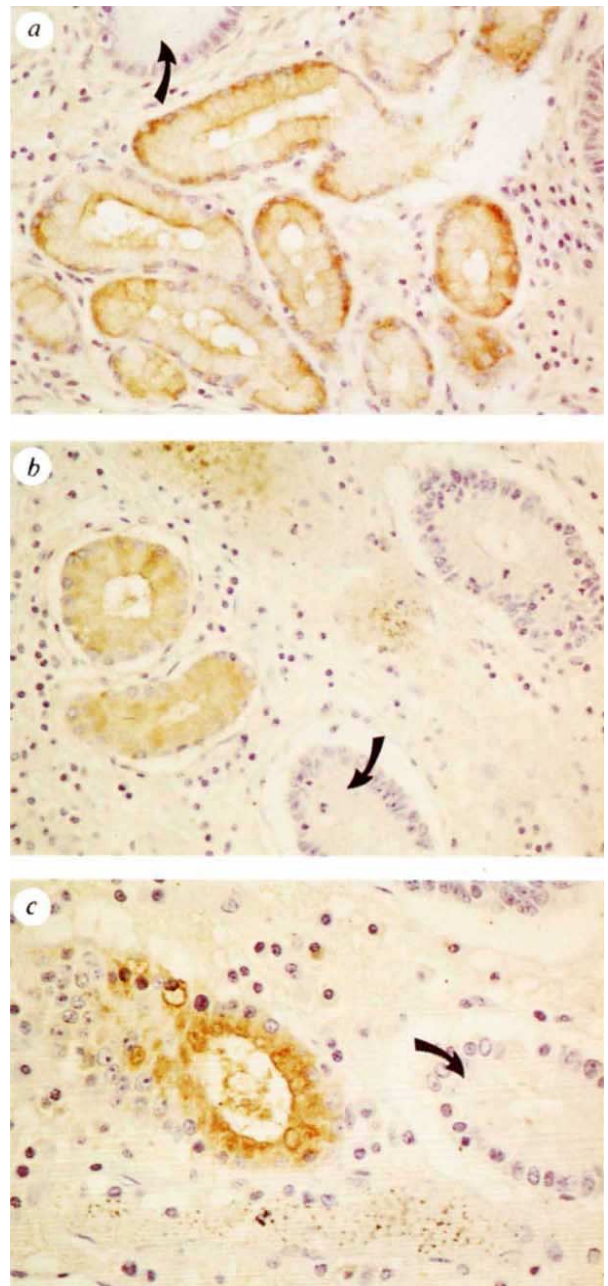


FIG. 4 The cell lineage in ileum (a, b) and colon (c) showing immunoreactive cytoplasmic and luminal EGF/URO. a, b, c, The normal crypt (arrow) shows negatively stained columnar cells and goblet cells. The staining is granular and varies considerably between individual cells in the lineage (a). The primary antibody was a polyclonal antibody against recombinant EGF/URO²⁷ raised in rabbits. The method was similar to that given in the legend to Fig. 3, apart from incubation in swine anti-rabbit antibody. Controls (absorption controls, absorbing the primary antibody with recombinant EGF/URO, and controls omitting the primary antibody) showed abolition of staining. Positive staining is also seen with a polyclonal antibody against recombinant human EGF/URO raised in sheep.

cell lineage stains uniformly negative with the monoclonal antibody Ki67, which recognizes a proliferation-dependent nuclear antigen in human cells¹⁶.

A search for other defined lineages of small intestinal origin in the newly formed glands reveals occasional Paneth cells (lysozyme-positive), a few chromogranin A and Grimelius-positive neuroendocrine cells and very infrequent but typical goblet cells, containing acid sialomucin.

These observations lead us to propose that mucosal ulceration is the signal for the development of this cell lineage. Development occurs close to the ulcer margin, indicating a concentration gradient of some inducing factor (possibly mesenchymal-derived¹³) with distance from the ulcer margin. It is initiated as a bud growing from the crypt base, where the intercalated multipotential undifferentiated stem cells are housed^{17,18}, and thus probably represents direct differentiation progeny of these stem cells. This bud pushes out into the lamina propria, forming a new gland. It is of interest that cell proliferation, as evidenced by lack of mitotic activity and Ki67 staining, is not involved in this morphogenesis; but the usually slowly cycling stem cells are fully capable of reducing their cell cycle time and increasing their cell production rate at times of need¹⁸. Moreover, because cells appear to migrate along the tubule, the local crypt stem cells would seem to be providing the entire lineage (although probably not the motive force for migration¹⁹). The stem cell zone hypothesis of Bjerkness and Cheng²⁰ states that below about cell position 5 in the crypt, cells migrate downwards towards the crypt base. This could indicate that the new lineage begins differentiation in this zone, and moves downwards to enter the newly formed tubule, as is further indicated by (1) the occasional occurrence of other progeny of stem cell differentiation (Paneth, endocrine and goblet cells) caught up in the downward migration, and (2) the occasional observation of cells similar to the lineage described within the bases of crypts without basal buds (presumably seen immediately before bud formation).

The tubules grow in the lamina propria, forming a new gland, and the cells acquire the ability to synthesize and secrete EGF/URO, which is carried to the surface by a tubule connecting the gland to the luminal surface of the villus. EGF/URO and neutral mucin are secreted into the lumen, and cells from the tubule migrate onto the villus surface, clothing its upper portion. The new glands are rarely found more than 1 cm from the ulcer margin, but as many as fifty individual glands can be found around small (5 mm) ileal ulcers; in larger ulcers, many more glands are found, and they can reach appreciable size (>5 mm diameter). They are commonest in the small intestine in Crohn's disease, present in all ulcers examined in 39 out of 45 cases, and although rarer, the same lineage is induced in the colon in Crohn's disease (5 out of 21 cases; Fig. 4). It is a constant finding in duodenal ulcer disease and duodenitis (36 out of 36 cases), and fairly common in gastric ulcer disease (15 out of 30 cases). Thus chronic mucosal ulceration anywhere in the gastrointestinal tract can induce this cell lineage from multipotential stem cells in intestinal crypts or gastric glands; the new glands so formed secrete EGF/URO.

EGF/URO, although an extremely potent stimulator of cell proliferation in the intestine when given parenterally^{7,8} is not absorbed and has no effect when administered lumenally when the mucosa is intact^{6,9,10}, even though there are now numerous reports of epidermal growth factor receptors on small intestinal and colonic epithelial cells²¹. It has instead been proposed that transforming growth factor alpha (TGF- α) could be the main ligand for EGF receptors in the gut²². But EGF/URO heals and prevents experimentally induced ulceration in the rat^{23,24}, and is absorbed through damaged mucosa²⁵; thus the EGF/URO produced locally by the new glands is available to regenerating epithelial and connective tissue cells in and around the ulcer. We therefore propose that a principal *in vivo* role of EGF/URO is to assist in the healing of ulceration in the gastrointestinal

mucosa through induction of this lineage.

What is the nature of this cell lineage? Similar proliferations have usually previously been regarded as metaplasias, called 'pyloric' or 'gastric' because of morphological resemblance to gastric epithelial cells, or even as Brunner's gland metaplasia on similar grounds²⁶. The latter could seem most reasonable because of the common EGF/URO production, but the lectin-binding profile and immunophenotype are different from those of both gastric and Brunner's gland epithelium, and seems unique among gastrointestinal epithelial cells. The term metaplasia implies a change from one defined differentiated cell type to another, but if, as these results indicate, this is a novel cell lineage, then it cannot strictly be regarded as a metaplasia. It could be that abnormal conditions induce novel cell lineages in epithelial stem cells, until now regarded as metaplasias, but only defined by detailed phenotyping. But it would be surprising if, generally speaking, such cells were never seen in normal conditions (J. Slack, personal communication). □

Received 3 October; accepted 7 November 1989.

- Gregory, H. *Nature* **257**, 325-327 (1975).
- Cohen, S. & Carpenter, G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1317-1321 (1975).
- Heitz, P., Kasper, M., Van Noorden, S., Polak, J. M., Gregory, H. & Pearce, A. G. E. *Gut* **19**, 408-413 (1978).
- Al-Nafussi, A. & Wright, N. *Virchows Arch. Cell Path.* **40**, 63-70 (1982).
- Malo, C. & Menard, D. *Gastroenterology* **83**, 28-35 (1982).
- Goodlad, R., Wilson, G., Lenton, W., Gregory, H., MacCullagh, K. & Wright, N. *Gut* **28**, 573-582 (1987).
- Thompson, T. S., Sharp, J. G., Saxena, S. K. & MacCullagh, K. *J. surg. Res.* **42**, 402-410 (1987).
- Walker-Smith, J. et al. *Lancet* **ii**, 1239-1240 (1985).
- Olsen, P. et al. *Gut* **25**, 1234-1240 (1984).
- Skov-Olsen, P., Poulsen, S. S., Kierkegaard, P. & Nexø, E. *Dig. Dis. Sci.* **29**, 615-622 (1984).
- Morson, B. & Dawson, I. *Gastrointestinal Pathology*, 293-312 (Blackwell, Oxford, 1979).
- Magnus, H. *Postgrad. Med. J.* **30**, 131-146 (1954).
- Richman, P. & Bodmer, W. *Int. J. Cancer* **39**, 317-328 (1987).
- Burchell, J. et al. *Cancer Res.* **47**, 5476-5482 (1987).
- Gendler, S., Taylor-Papadimitriou, J., Duhig, T., Rothbard, J. & Burchell, J. *J. biol. Chem.* **263**, 12820-12823 (1988).
- Schwartz, R. et al. *J. Immunol. Meth.* **90**, 365-371 (1986).
- Cheng, H. & Leblond, C. P. *Am. J. Anat.* **141**, 537-562 (1974).
- Wright, N. & Alison, M. R. *The Biology of Epithelial Cell Populations Vol. 2* (Clarendon, Oxford, 1984).
- Kaur, P. & Potten, C. S. *Cell Tissue Kinet.* **19**, 601-610 (1986).
- Bjerknes, R. & Cheng, H. *Am. J. Anat.* **160**, 76-92 (1981).
- Thompson, J. *Am. J. Physiol.* **264**, G429-G435 (1988).
- Barnard, J. et al. *Gastroenterology* **96**, A27 (1989).
- Kierkegaard, P., Olsen, P. S., Poulsen, S. S. & Nexø, E. *Gastroenterology* **85**, 1277-1283 (1983).
- Konturek, S., Dembinski, A., Warzecha, A., Brzozowski, T. & Gregory, H. *Gastroenterology* **94**, 1300-1307 (1988).
- Poulsen, S. *Scand. J. Gastroenterol. (Suppl. 22)* **128**, 20-21 (1987).
- Lee, F. *J. pathol. Bact.* **87**, 267-277 (1964).
- Smith, J. et al. *Nucleic Acids Res.* **10**, 4467-4482 (1982).

ACKNOWLEDGEMENTS. We thank Gordon Stamp, Paul Richman and Jonathan Slack for discussions, Harry Gregory for the gift of the anti-human EGF/URO antibodies, Joyce Taylor-Papadimitriou for SM3 and HMFG-1 and -2 antibodies, Paul Richman for the 3B10 antibody, and Chris Foster for lectins.

High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients

Shen Pang, Yoshio Koyanagi, Steven Miles, Clayton Wiley*, Harry V. Vinters† & Irvin S. Y. Chen‡

* Department of Pathology, UCSD School of Medicine, La Jolla, California 92083, USA

† Departments of Pathology†, and Microbiology and Immunology and Division of Hematology-Oncology, Department of Medicine‡, UCLA School of Medicine and Jonsson Comprehensive Cancer Center, Los Angeles, California 90024-1678, USA

IN the host cell, retroviral DNAs exist in three main forms: unintegrated linear, unintegrated circular, and integrated (the provirus)¹. High levels of unintegrated forms of retroviral DNA often correlate with superinfection and accompanying cytopathic effects²⁻⁸, as, for example, in the case of feline acquired immunodeficiency^{7,8}. In culture, HIV-1 infection also results in high levels of unintegrated viral DNA although direct correlations with cytopathicity have not been made^{9,10}. The low frequency of

HIV-1-infected cells in patients^{11,12} has made it difficult to determine the structure of the viral DNA in fresh tissue samples from AIDS patients by standard methods such as Southern hybridization. The PCR technique^{13,14}, however, which allows the detection of viral DNA at levels far below that possible by other hybridization methods is, in its conventional form, of limited use for quantitative analysis. To study the amount and form of HIV-1 DNA in primary tissue of AIDS patients we have therefore modified the PCR method. Our results indicate that each of the three species of viral DNA are detectable in blood and brain of AIDS patients and that in autopsy samples from patients with HIV encephalitis there is a considerably higher proportion of unintegrated viral DNA.

We adapted the PCR method, which in its standard form is not easily used in a quantitative fashion by having one oligonucleotide of each pair labelled at the 5' end with $\gamma^{32}\text{P}$ -ATP. The ^{32}P -labelled PCR products then obtained by amplification were visualized directly by acrylamide gel electrophoresis and autoradiography (Fig. 1b), and quantitated by excision of bands and scintillation counting. We found that twofold differences could be easily resolved which was sufficient for the application described here. Details of the quantitative procedure can be obtained on request³³.

Oligonucleotide probes were chosen from conserved regions of HIV-1 (Fig. 1a), and tested for specificity on three cloned

isolates (HIV-1_{JR-CSF}^{15,16}, HIV-1_{SF2}, and HIV-1_{NL}). In most experiments, we used LTR-gag oligonucleotides (Fig. 1a, M667 and M661), as PCR products from this region gave the clearest resolution. Additionally, we anticipated that these probes would detect completely or almost completely synthesized viral DNA rather than incomplete reverse transcriptase products because, according to current models for reverse transcription of retroviral RNA^{17,18}, the region spanning gag and LTR sequences should be one of the last synthesized.

Table 1 shows levels of viral DNA present in brain and blood cells. Viral DNA is detected at levels as low as 25 molecules per 10⁵ cells. To quantitate the relative amounts of total cellular DNA, we used labelled oligonucleotides homologous to β -globin as internal PCR standards^{13,14}. In general, the relative abundance of viral DNA in total brain-cell DNA from HIV encephalitis autopsy samples was significantly greater than that detected in peripheral blood-cell DNA from other AIDS patients or autopsy brain samples of non-HIV encephalitis patients (Table 1), consistent with rough estimates of virus infection made by other methods^{9,19}. Occasional blood samples (from C.D.) showed higher levels of HIV-1 DNA.

Integrated and unintegrated species of viral DNA were separated by fractionation of total cell DNA into high and low molecular mass portions by agarose gel electrophoresis in 'low melting point' agarose. The high and low molecular mass DNA

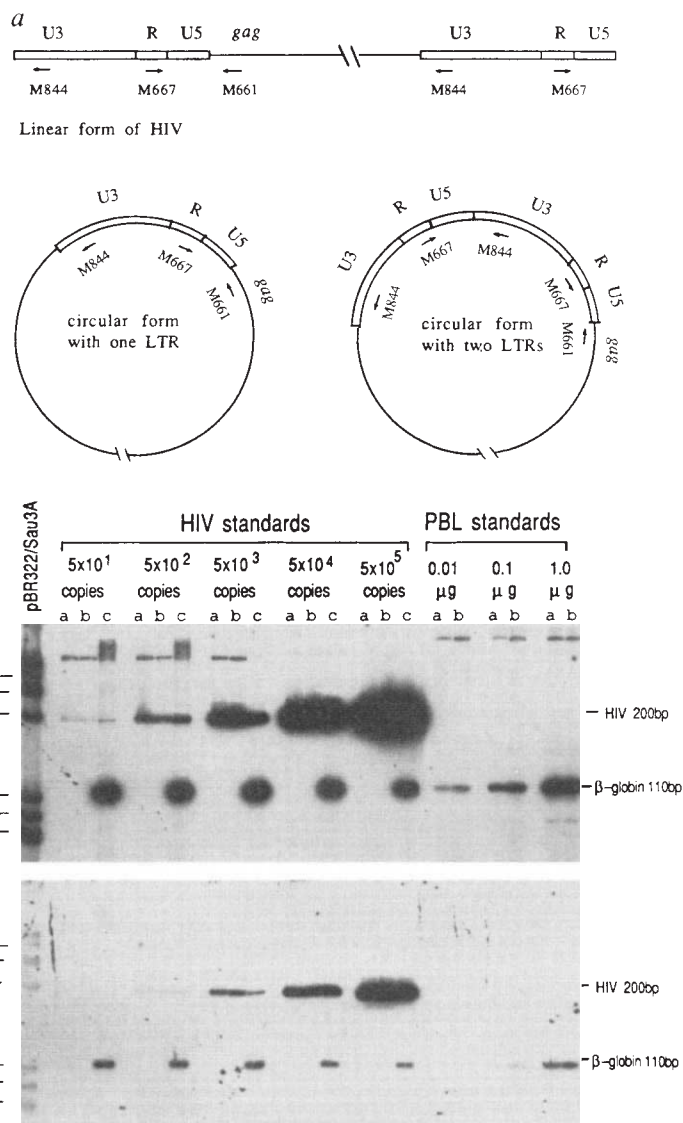


FIG. 1 PCR analysis of HIV-1 *in vivo*. a, Primers and probes for polymerase chain reactions (PCR). The nucleotide (nt) sequence of M667 is GGCTAACTAGGGAACCCACTG and positioned at 496–516 on the sequence of HIV-1_{SF2}(ARV-2)³¹; the sequence of M661 is CTGCGTCGAGAGAGCTCCTCTGG positioned at 696–673; the sequence of M844 is CTGATCCCTGGCCTGGTGTG positioned at 113–93. In the experiments, the 5' end of M667 was labelled with ^{32}P . b, Standardization of genomic DNA and HIV-1 DNA by PCR. DNA (0.01–1 μg) isolated from normal PBL and cloned HIV-1_{JR-CSF} DNA equivalent to 50–500,000 copies of HIV-1 genomes were tested by PCR analysis. In the HIV-1 standards, the first lanes (a and b) of each group contain 0.01 μg λ DNA and cloned HIV-1 DNA as indicated, and the third lane (c) contains an additional 1 μg uninfected human PBL DNA. As the cloned HIV-1 DNA is supercoiled, and these forms are resistant to heat denaturation, digestion with *Sph*I was performed before PCR analysis. The predicted size of the amplified region from HIV-1_{JR-CSF} is 200 base pairs (bp) and the predicted size of the β -globin amplified fragment is 110 bp^{13,14}. Plasmid pBR322 digested with *Sau* 3AI serves as a molecular mass marker. The lower panel is a shorter exposure of the same autoradiograph.

METHODS. Brain samples were obtained within 24 h of the death of patients, and stored at -70°C . Tissue (~ 0.3 g) was homogenized in liquid nitrogen, lysed by Hirt buffer (0.5% SDS 10 mM Tris-HCl buffer pH 7.4, and 1 mM EDTA), and digested with proteinase K (250 μg/ml final concentration) overnight at 37°C . Samples were purified by phenol and chloroform/isoamyl alcohol extraction. Mononuclear cells were isolated from blood samples by Ficoll-Hypaque gradient centrifugation. The purification of total nucleic acid was similar to that of brain tissue DNA. Details of the quantitative PCR method are available on request³³. Oligonucleotides M667 and M661 were used to detect HIV-1 DNA, and PCO3 and PCO4 were used to detect the β -globin gene^{13,14}. M667 and PCO3 were labelled with ^{32}P on the 5' end. The band derived from the β -globin gene is 110 bp, and the HIV-1 bands are 200 bp.

TABLE 1 Amount of unintegrated and integrated viral DNA in brain and blood

Patient	Tissue	Other neuropathology	Fraction	Repeat	HIV-1 copy Number per 10 ⁴ cells	Unintegrated/ integrated
HIV encephalitis						
C.T.	Brain	CMV encephalitis	total	3	281	11
			high <i>M_r</i>	3	27	
			low <i>M_r</i>	3	287	
J.C.	Brain	Foci of white matter necrosis	total	5	124	12
			high <i>M_r</i>	3	8	
			low <i>M_r</i>	3	92	
J.G.(1)	Brain	Cerebral lymphoma	total	3	33	11
J.G.(2)	Brain		total	3	172	
			high <i>M_r</i>	3	9	
			low <i>M_r</i>	3	99	
J.P.(1)	Brain		total	3	37	
J.P.(2)	Brain		total	4	81	
			high <i>M_r</i>	3	1	
J.R.	Brain		CMV encephalitis, diffuse necrotizing leukoencephalopathy	low <i>M_r</i>	3	81
		total		5	186	
		high <i>M_r</i>		3	5	
K.N.	Brain	Early vacuolar myelopathy	low <i>M_r</i>	3	104	19
			total	3	139	
			high <i>M_r</i>	2	4	
A86-42	Brain	CMV encephalitis	low <i>M_r</i>	2	78	6
			total	2	122	
			high <i>M_r</i>	2	16	
A86-90	Brain	CMV encephalitis	low <i>M_r</i>	2	100	14
			total	2	367	
			high <i>M_r</i>	2	10	
JR 1013 (paediatric)	Brain	Inflammation of the ependymal region, possible adenovirus encephalitis	low <i>M_r</i>	2	143	9
			total	1	58	
			high <i>M_r</i>	1	9	
J.C.	Blood		low <i>M_r</i>	1	85	0.2
			total	3	11	
			high <i>M_r</i>	2	6	
			low <i>M_r</i>	2	1	
Non-HIV encephalitis						
A86-16	Brain	Normal	total	2	14.9	0.1
			high <i>M_r</i>	2	6.8	
			low <i>M_r</i>	2	1.0	
A86-103	Brain	Normal	total	2	1.5	0.8
			high <i>M_r</i>	2	1.9	
			low <i>M_r</i>	2	1.5	
A87-18	Brain	CMV encephalitis	total	1	<1	
			high <i>M_r</i>	1	<1	
			low <i>M_r</i>	1	<1	
A87-45	Brain	CMV encephalitis	total	1	<1	
			high <i>M_r</i>	1	<1	
			low <i>M_r</i>	1	<1	
R.S.	Brain	Focal subependymal gliosis	total	1	2.5	1.4
			high <i>M_r</i>	1	1	
			low <i>M_r</i>	1	1.4	
E.B.	Brain	Normal	total	2	52	0.3
			high <i>M_r</i>	2	55	
			low <i>M_r</i>	2	15	
P.K.	Brain	Normal	total	2	11	0.2
			high <i>M_r</i>	2	10	
			low <i>M_r</i>	2	2	
C.D.	Blood		total	2	20	11
			high <i>M_r</i>	1	1	
			low <i>M_r</i>	1	11	
G.J.	Blood		total	3	33	10
			high <i>M_r</i>	3	5	
			low <i>M_r</i>	3	50	
G.P.	Blood		total	2	5	0.3
			high <i>M_r</i>	2	1	
			low <i>M_r</i>	2	0.4	
M.T.	Blood		high <i>M_r</i>	2	2	9
			low <i>M_r</i>	2	19	
			total	2	10	
R.C.	Blood		high <i>M_r</i>	4	5	2
			low <i>M_r</i>	4	8	
			total	4	10	
V.C.	Blood		high <i>M_r</i>	4	6	2
			low <i>M_r</i>	4	13	
			total	2	2	
E.B.	Blood		high <i>M_r</i>	2	4	1
			low <i>M_r</i>	2	4	
			total	2	4	
P.K.	Blood		total	2	4	0.3
			high <i>M_r</i>	2	4	
			low <i>M_r</i>	2	1	

In most cases, the calculated amounts are the average of two to five independent repeat experiments. For each experiment the PCR products were subjected to gel electrophoresis and visualized by autoradiography. The bands specific to HIV-1 and β -globin were excised and quantitated by liquid scintillation counting. Standard curves were generated from the standards, and the number of HIV-1 copies calculated for each sample. The amount of globin-specific signal in each total or high molecular mass ('high' in table) DNA sample allowed normalization for amounts of total DNA. The unintegrated HIV-1 DNA was further corrected for any degradation or shearing of genomic DNA containing integrated HIV-1 DNA by subtracting the calculated 'contamination value' determined by the intensity of the β -globin signals and the level of integrated HIV-1. Theoretically, the globin signal in low molecular mass DNA fractions ('low' in table) should be zero; however, some globin signal is detected (see Fig. 2) due to degradation or shearing of the DNA during preparation. Two determinations are shown for total brain DNA of J.G. and J.P. because different brain tissue fragments were used for the DNA isolation and variable amounts of viral DNA are present in different portions of a sample. The differences in both the ratios and copy number of viral DNAs in HIV encephalitis brain tissue versus blood are highly significant, with $P < 0.01$ and $P < 0.005$, respectively (T -test)³². A similar result was observed by comparing the HIV encephalitis samples with the non-HIV encephalitis samples with $P < 0.05$ (T -test).

fractions containing chromosomal DNA plus integrated viral DNA and low molecular mass cellular DNA plus unintegrated viral DNA, respectively, were subjected to PCR, along with HIV-1 and cellular DNA β -globin standards (Fig. 2). As the complexity of the DNA sample (Fig. 1b) or the amount of total DNA in the reaction has no effect on detection of HIV-1 in the samples, quantitative measurements of the relative HIV-1 copy number in each fraction are possible. Complete separation of high and low molecular mass DNAs should result in all of the β -globin signal within the high molecular mass DNA fractions; however, due to some degradation or shearing, or both, of the DNAs, either at autopsy or during preparation, some β -globin signal is also seen in the low molecular mass fractions. The viral DNA from low molecular mass DNA fractions therefore includes integrated viral DNA in proportion to the amount of β -globin detected in that fraction. This low extent of HIV-1 signal was controlled for by subtracting the amount of integrated viral DNA, calculated from the intensity of the β -globin PCR signal, from the amount of viral DNA present in a low molecular mass fraction (Table 1).

Our results show that both integrated and unintegrated viral DNA were detected in all tissues (Table 1). Examples of the fractionation are shown in Fig. 2. A comparison of the relative amounts of integrated and unintegrated viral DNA in brain and blood, indicates that there is consistently more unintegrated viral DNA than integrated viral DNA in brain tissue from HIV encephalitis cases. The ratio of unintegrated to integrated viral DNA ranged from 6-fold to more than 80-fold in these brain tissues. From blood samples as well as brain tissue from non-HIV encephalitis patients, the ratio of unintegrated to integrated viral DNA was variable, although both species were consistently detected. In one HIV encephalitis case, where we were able to obtain blood and brain samples from the same individual (J.C.),

a similar pattern showing a greater proportion of unintegrated DNA in brain samples was observed (Fig. 2b).

We examined the form of the unintegrated viral DNA by choosing PCR oligonucleotide primers that would discriminate between circular forms of viral DNA having two LTRs and other unintegrated viral DNA species (Fig. 1a). Primers that detect circular viral DNA were chosen such that they would cross the 'circle junction' between U3 and U5 of two LTRs. Two LTR circular forms of viral DNA were detectable in brain tissue at a low level, but far less than the amount of viral DNA detected using LTR-*gag* primers (Fig. 3). As both primers amplify DNA efficiently in other PCR applications, the circular form can, assuming equal efficiency of PCR, be estimated to be ~0.1–1.0% of the total HIV-1 DNA. Thus, most unintegrated viral DNA is not of the two-LTR circular form.

Most cells in the brain of HIV encephalitis patients identified as expressing HIV-1 by immunohistochemical and *in situ* hybridization methods have been of monocyte/macrophage or microglial phenotypes^{9,20–23}. Given the high levels of unintegrated viral DNA, it is possible that the replication of HIV-1 in these cells differs from that in T cells^{24,25}, which may be the predominant type infected in the blood. Initial experiments examining the structure of viral DNA in blood monocytes did not show major differences from that of the T cells (data not shown); this may, however, reflect the differentiation state or micro-environment of the blood monocytes as opposed to tissue macrophages present in the brain.

Unintegrated viral DNA is generally associated, *in vitro*, with superinfection and re-infection, where viral DNA is generated by reverse transcription of newly infecting virions^{2–8}. We tested the filtrate of homogenized brain tissue for presence of viral antigens. In some cases (J.C., J.R., C.T. and J.G.), high levels of viral *gag* p24 antigen were detected; in other cases (K.N.

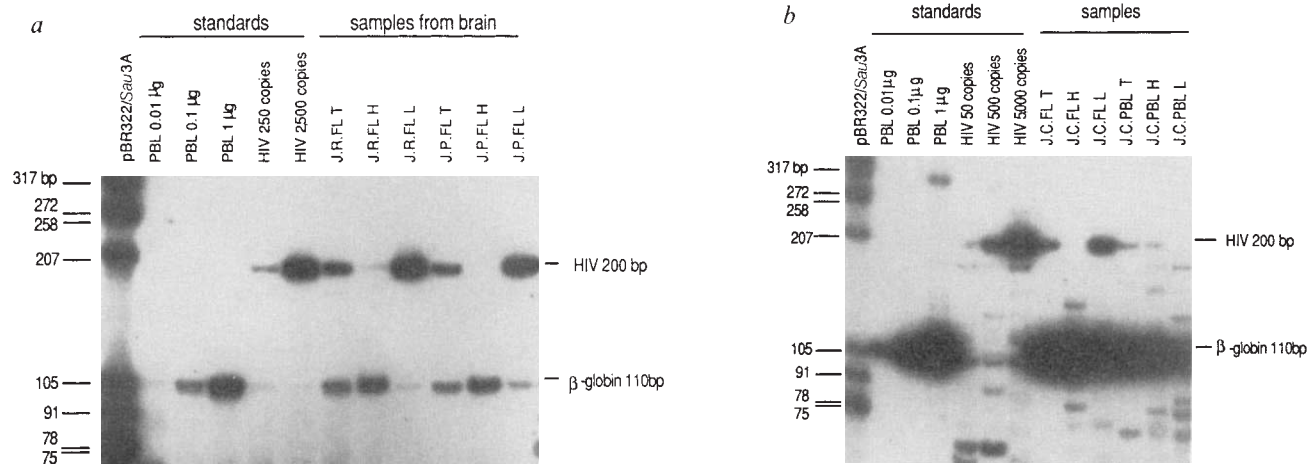
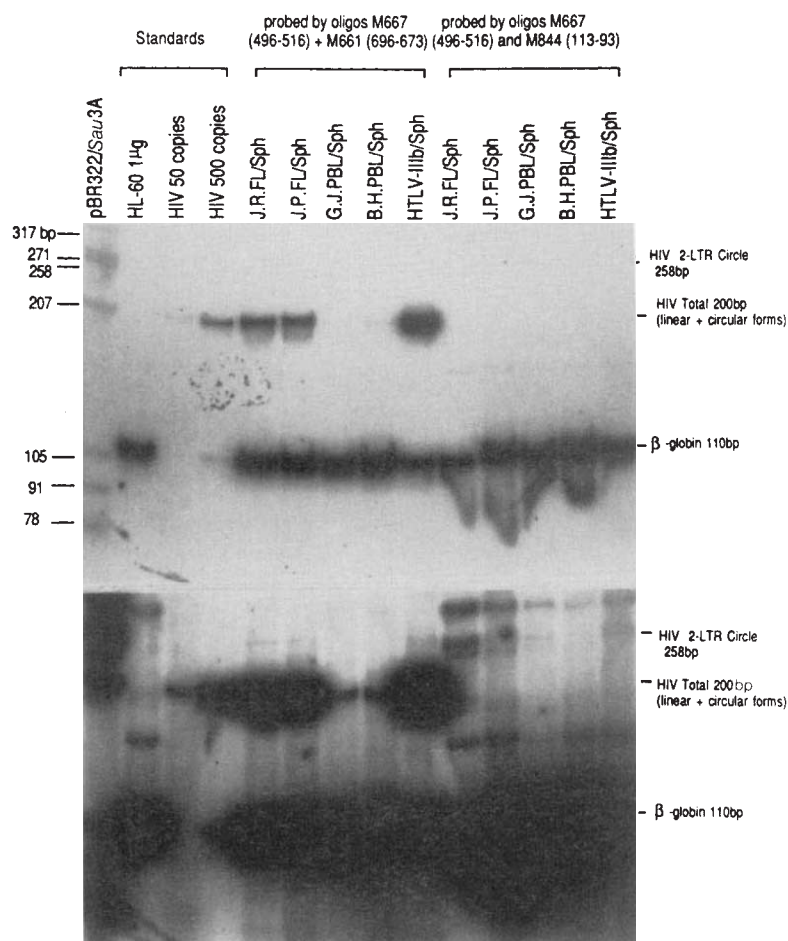


FIG. 2 Fractionation and detection of HIV-1 high and low molecular mass DNA. *a*, HIV-1 DNA isolated from brains of autopsy cases. PBL DNA samples (0.01 μ g, 0.1 μ g and 1.0 μ g) isolated from normal blood donors served as β -globin gene standards and negative controls for HIV-1 DNA. Cloned HIV-1_{JR-CSF} DNA plus carrier PBL DNA served as the viral DNA standard. Total DNAs (T) from patient samples were separated into high (H) and low (L) molecular size fractions by 'low melting point' agarose gel electrophoresis. Abbreviations: J.R.FL T, total DNA from J.R. brain (frontal lobe), 0.1 μ g; J.R.FL H, J.R. high molecular mass DNA; and J.R.FL L, low molecular mass DNA from patient J.R. Samples for lanes J.R.FL H and J.R.FL L were derived from 1 μ g total J.R.FL DNA and fractionated and analysed by PCR. Samples, J.P.FL T, J.P.FL H, and J.P.FL L were assayed similarly. The β -globin standards were run in the same PCR to estimate DNA recovery and the extent of DNA shearing or degradation. Plasmid pBR322 was digested with *Sau* 3AI to serve as the molecular mass marker. *b*, HIV-1 DNA in brain of J.C. (J.C.FL) and blood (J.C.PBL). The total DNA lanes of J.C.FL T and J.C.PBL T contained 0.5 μ g DNA. J.C.FL H and J.C.FL L were derived from 1 μ g J.C.FL DNA, and

J.C.PBL L and J.C.PBL H were derived from 1 μ g of J.C.PBL DNA. The indicated amounts of PBL, cloned HIV-1_{JR-CSF} DNA and DNA from patient J.C. were assayed as in Fig. 2a.

METHODS. The purified DNA was loaded onto a 0.8% low melting point agarose gel which was run in 0.5 $\times$ E buffer (20 mM Tris-Cl buffer pH 7.4, 10 mM NaAc and 1 mM EDTA) at 2 V cm^{-1} overnight, with ethidium bromide dye at 0.5 $\mu\text{g ml}^{-1}$. DNA bands were collected under ultraviolet illumination. In most cases, the high molecular mass genomic DNA migrated very slowly, and showed a strong band close to the origin of the gel. In some cases (J.C.PBL), the genomic DNAs were partially sheared and formed a wider band. In these, DNAs larger than 15 kilobases (kb) were considered to be of high molecular mass, whereas those smaller than 15 kb but larger than 4 kb were considered to be of low molecular mass. The agarose-containing fractionated DNAs were melted at 65 $^{\circ}\text{C}$, followed by extraction with phenol and chloroform:isoamyl alcohol. The DNA samples underwent 25 cycles of PCR amplification, as described in the legend of Fig. 1, with 5×10^5 c.p.m. of ^{32}P -labelled PCO3 added to each sample (Fig. 1b).

FIG. 3 Detection of circular viral DNA in brain and blood. HL-60 DNA (1 µg) served as a negative control for HIV-1 sequences, and the indicated amounts of HIV-1_{JR-CSF} served as standards. The middle set of lanes probed by oligonucleotides M667 and M661 (LTR-*gag*) indicate the total amount of viral DNA. The right set of lanes demonstrates detection of circular forms of viral DNA with two LTRs of the same DNA samples using oligonucleotides M667 and M844 (see Fig. 1a). The two sets of DNA samples underwent PCR analysis in parallel with the different primers indicated. In addition, the oligonucleotides that are homologous to the β-globin gene were added to all samples. Both pairs of HIV-1-specific oligonucleotides were added to the negative control. Only M667 and M661 oligonucleotides were added to HIV-1 DNA standards. The PCR was performed for 30 cycles, according to the methods in Fig. 1. The lower panel is a longer exposure of the same autoradiograph. As circular forms of HIV-1 DNA tend to denature incompletely during PCR, the samples were digested with *Sph* I (an enzyme site present in a conserved region of *gag* of all sequenced HIV-1 isolates) to linearize circular forms of viral DNA before the PCR. Two bands are detected, with sizes of about 250 bp. Both are HIV-1-specific as neither are evident in the negative controls. The top band of length 258 bp matches the size calculated from HIV-1 sequence data. The lower band is ~10 bp shorter, possibly due to deletion of sequence during the PCR.



and J.P.), no viral antigen was detected although viral DNA was evident in the same sample (see Table 1). Thus, there seemed to be no correlation between levels of unintegrated viral DNA and presence of viral antigens which indicates that it may be possible for unintegrated viral DNA in the brain tissue to result from processes independent of virion infection or re-infection. These may involve persistence, or a latent state of the virus in an unintegrated form, or intracellular reverse transcription of viral DNA in the absence of virion production. Analysis of viral RNA in the cells may help to distinguish between these two possibilities.

High levels of unintegrated viral DNA are also associated with infection by other lentiviruses such as visna virus²⁶⁻²⁸. Studies with avian leukaemia viruses (ALV) and reticulo-endotheliosis viruses *in vitro*, and some strains of feline leukaemia viruses (FeLV) both *in vitro* and in infected cats, demonstrate a close correlation between cytopathic effects of the virus and the presence of high levels of unintegrated viral DNA^{2-8,29,30}. Our results demonstrate a correlation between HIV encephalitis and the presence of unintegrated viral DNA in brain tissue. It would be of interest to determine whether high levels of unintegrated viral DNA correlate with the onset of pathogenic effects of HIV-1 in brain of patients with HIV encephalitis or blood of patients at different stages of disease. Such studies will require a larger scale longitudinal study of HIV-1-infected individuals. □

3. Weller, S. K., Joy, A. E. & Temin, H. M. *J. Virol.* **33**, 494-506 (1980).
4. Weller, S. K. & Temin, H. M. *J. Virol.* **39**, 713-721 (1981).
5. Chen, I. S. Y. & Temin, H. M. *J. Virol.* **41**, 183-191 (1982).
6. Branic, M., Stowring, L., Ventura, P. & Haase, A. T. *Nature* **292**, 240-242 (1981).
7. Hoover, E. A., Mullins, J. I., Quackenbush, S. L. & Gasper, P. W. *Blood* **70**, 1880-1892 (1987).
8. Mullins, J. I., Chen, C. S. & Hoover, E. A. *Nature* **319**, 333-336 (1986).
9. Shaw, G. M. *et al. Science* **227**, 177-181 (1985).
10. Ho, D. D. *et al. New Engl. J. Med.* **313**, 1493-1497 (1985).
11. Harper, M. H., Marselle, L. M., Gallo, R. C. & Wong-Staal, F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 772-776 (1985).
12. Shaw, G. M. *et al. Science* **226**, 1165-1170 (1984).
13. Saiki, R. K. *et al. Science* **230**, 1350-1354 (1985).
14. Saiki, R. K. *et al. Science* **239**, 487-491 (1988).
15. Koyanagi, Y. *et al. Science* **236**, 819-822 (1987).
16. Cann, A. J., Koyanagi, Y. & Chen, I. S. Y., in *Control of Human Retrovirus Gene Expression* (eds Franza, R., Cullen, B. & Wong-Staal, F.) 127-133 (Cold Spring Harbor Laboratory, New York, 1988).
17. Varmus, H. & Swanstrom, R. in *RNA Tumor Viruses* (eds Weiss, R., Teich, N., Varmus, H. & Coffin, J.) 369-512 (Cold Spring Harbor Laboratory, New York, 1984).
18. Varmus, H. & Swanstrom, R. in *RNA Tumor Viruses* (eds Weiss, R., Teich, N., Varmus, H. & Coffin, J.) 75-134 (Cold Spring Harbor Laboratory, New York, 1985).
19. Harper, M. E. *et al. New Engl. J. Med.* **315**, 1073-1078 (1986).
20. Price, R. W. *et al. Science* **239**, 586-592 (1988).
21. Srinivasan, A., Dorsett, D., York, D., Bohan, C. & Anand, R. *Arch. Virol.* **99**, 135-141 (1988).
22. Koenig, S. *et al. Science* **223**, 1089-1093 (1986).
23. Wiley, C. A., Schrier, R. D., Nelson, J. A., Lampert, P. W. & Oldstone, M. B. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7089-7093 (1986).
24. Anand, R., Srinivasan, A., Gardner, M. B., Luciw, P. A. & Dandekar, S. *Ann. Neurol.* **23** (Suppl.), S62-S65 (1988).
25. Fauci, A. S. *Science* **239**, 617-622 (1988).
26. Branic, M., Haase, A. T. *Proc. natn. Acad. Sci., U.S.A.* **75**, 6125-6129 (1978).
27. Harris, J. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 7212-7215 (1984).
28. Haase, A. T., Stowring, L., Narayan, O., Griffin, D. & Price, D. *Science* **195**, 175-177 (1977).
29. Overbaugh, J., Donahue, P. R., Quackenbush, S. L., Hoover, E. A. & Mullins, J. I. *Science* **239**, 906-910 (1988).
30. Overbaugh, J., Riedel, N., Hoover, E. A. & Mullins, J. I. *Nature* **332**, 731-734 (1988).
31. Sanchez-Pescador, R. *et al. Science* **227**, 484-492 (1985).
32. *SAS User's Guide: Statistics Version 5th ED.* (SAS Institute, 1985).
33. Requests for supplementary data should be addressed to the London editorial office of *Nature*.

Received 24 August; accepted 7 November 1989.

1. Varmus, H. *Science* **240**, 1427-1435 (1988).
2. Keshet, E., O'Rear, J. J. & Temin, H. M. *Cell* **16**, 51-61 (1979).

Inhibition of HIV-1 protease in infected T-lymphocytes by synthetic peptide analogues

Thomas D. Meek*, Dennis M. Lambert†, Geoffrey B. Dreyer*, Thomas J. Carr*, Thaddeus A. Tomaszek, Jr*, Michael L. Moore‡, James E. Strickler§, Christine Debouck||, Lawrence J. Hyland*, Thomas J. Matthews¶, Brian W. Metcalf* & Stephen R. Petteway†

Departments of * Medicinal Chemistry, † Anti-Infectives, ‡ Peptide Chemistry, § Macromolecular Sciences and || Molecular Genetics, Smith Kline & French Laboratories, 709 Swedeland Road, King of Prussia, Pennsylvania 19406, USA
¶ Department of Surgery, Duke University Medical Center, Durham, North Carolina 27710, USA

THE *gag* and *pol* genes of the human immunodeficiency virus type 1 (HIV-1) (ref. 1) are translated as two polypeptides, Pr55^{gag} and Pr160^{gag-pol} (refs 2–6), which are subsequently cleaved by the action of a virus-encoded protease into the four structural *gag* proteins of the virion core (p17, p24, p7 and p6)^{7–9} and the *pol*-encoded enzymes essential for retrovirus replication (protease, reverse transcriptase, ribonuclease H, and endonuclease). Mutational inactivation of the proteases of HIV-1^{10–13} and other retroviruses^{14,15} results in immature, non-infectious virions, indicating that exogenous inhibition of the protease may represent an attractive approach to anti-AIDS therapy. Here we demonstrate that synthetic peptide analogues, which are potent inhibitors of purified HIV-1 protease, inhibit the processing of the viral polypeptides in cultures of HIV-1-infected T lymphocytes and attenuate viral infectivity.

The expression of HIV-1 protease in *Escherichia coli*¹⁶ and its subsequent characterization showed that it belongs to the mechanistic class of the aspartic proteases¹⁷, and that in its active form the enzyme is a homodimer^{17–20}. Potent inhibitors of other aspartic proteases such as porcine pepsin²¹ and human renin²² have been developed in which the scissile dipeptide of an oligopeptide substrate is replaced by a hydroxyethylene-containing isosteric moiety. Similarly, we have shown that the substitution of the scissile Tyr-Pro dipeptide of substrates such as Ac-Ser-Ala-Ala-Tyr-Pro-Val-Val-NH₂ by a hydroxyethylene isostere analogue of the dipeptide Phe-Gly, (4*S*, 5*S*)-4-hydroxy-5-amino-6-phenylhexanoic acid (Pheψ[CH(OH)CH₂]Gly), affords competitive inhibitors of HIV-1 protease which exhibit inhibition constants 4–5 orders of magnitude lower than the Michaelis constants of their corresponding oligopeptide substrates²³. Four inhibitors of this type, kinetically characterized as previously described^{23,24}, were evaluated for their ability to inhibit HIV-1 protease within a culture of T-lymphocytes infected with HIV-1. These compounds and their inhibition constants measured with purified HIV-1 protease are as follows: compound 1, Cbz-Ala-Pheψ[CH(OH)CH₂]Gly-Val-Val-OCH₃ (*K*_i = 120 nM); compound 2, Ala-Ala-Pheψ[CH(OH)CH₂]Gly-Val-Val-OCH₃ (*K*_i = 18 nM)²³; compound 3, Cbz-Ala-Ala-Pheψ[CH(OH)CH₂]Gly-Val-Val-OCH₃ (*K*_i = 48 nM); and compound 4, Boc-Ser-Ala-Ala-Tyrψ[CH(OH)CH₂]Gly-Val-Val-OCH₃ (*K*_i = 180 nM)²³ (Cbz, benzyloxycarbonyl; Boc, *t*-butyloxy-carbonyl). These inhibitors are equivalent in size or smaller than typical oligopeptide substrates of the protease^{17,23–25}, and are considerably more potent against HIV-1 protease than is pepstatin A (*K*_i = 1.4 μM)²³, a generic peptide analogue inhibitor of the aspartic proteases.

We treated T lymphocytes chronically infected with HIV-1 (CEM/III_B cells²⁶) with compounds 1, 3 and pepstatin A as described in Fig. 1. Western blot analysis of cell lysates with

monoclonal antibodies reactive with Pr55^{gag}/p24 (Fig. 1*a*) or Pr55^{gag}/p17 (Fig. 1*b*) demonstrated that proteins corresponding to Pr55^{gag}, p24 and p17 were detectable in the control cultures (lane C). Pepstatin A had no apparent effect on the relative proportions of these *gag* proteins. By contrast, when cultures were exposed to compounds 1 and 3, intermediate products of relative molecular mass (*M*_r) 40,000 (40K) and 47K appeared, and in the case of compound 1, an intermediate of *M*_r 25K was also detectable. These intermediates are probably the p17–p24, p17–p24–(p1)–p7, and p24–(p1) proteins, respectively, which would result from inhibition of cleavages within Pr55^{gag} at amino acids 377–378, 448–449 and 363–364^{2,3,7–9,25,27}. Furthermore, there was a relative decrease, especially with compound 3, in the appearance of p24 when compared with control or pepstatin-treated cultures, which is noteworthy considering the continuous production of virus in the chronically infected cells. The monoclonal antibody against p17 (Fig. 1*b*) identified the same two intermediates as that specific for p24 (compound 3 (lane 3) and compound 1 (lane 1)). There was also a decrease, especially with compound 3, in relative levels of p17 when compared with the untreated control. Neither antibody detected any of the 25K, 40K, or 47K processing intermediates in samples treated with compound 2 (20 μM), pepstatin (20 μM or 100 μM) or compound 4 (100 μM). Results identical to those shown in Fig. 1*a*, *b* were obtained for compounds 1 and 3 when they were added to cell cultures in a single concentration of 20 μM for 24 h (data not shown).

It was anticipated that inhibition of HIV-1 protease activity in cell culture would similarly affect processing of reverse transcriptase from Pr160^{gag-pol} to its 66K and 51K subunits. The CEM/III_B cell extracts were examined by western blot analysis using murine monoclonal antibody specific for HIV-1 reverse transcriptase (Fig. 1*c*). Treatment with either compound 1 or 3 caused a marked reduction in detectable 66K and 51K bands for reverse transcriptase and an increased accumulation of *gag-pol* processing intermediates. Cultures treated with pepstatin A show little or no effect compared with untreated controls. These results clearly demonstrate that compounds 1 and 3 inhibit the processing of Pr160^{gag-pol} to its requisite subunit components.

Seelmeier *et al.*¹³ reported that treatment of HIV-infected H9 cells with pepstatin A resulted in partial inhibition of Pr55^{gag} processing, as shown by the diminished conversion of Pr55^{gag} to a 40K intermediate. A decrease in the levels of p17 and p24 was not detected in the pepstatin-treated cells or virions, an anticipated result, given the large steady-state concentrations of mature virion proteins within the infected cell culture. To analyse the effects of inhibitor treatment on processing of Pr55^{gag} synthesized *de novo* in chronically infected cells, we labelled the CEM/III_B cells with [³⁵S]methionine following a 1 h pre-exposure to 100 μM of the inhibitors (Fig. 2). After initial pulse-labelling, immunoprecipitated proteins corresponding to *env* gp160 and Pr55^{gag} were detected (Fig. 2, control 0). Following subsequent chase, increasing amounts of gp120 and p24 were observed at 1.5 and 3 h (Fig. 2, controls 1.5 and 3). In the presence of compounds 1–3 at 100 μM, inhibition of the processing of Pr55^{gag} was apparently complete even after 3 h of chase (Fig. 2). The inactivity of compound 2 observed in the western blot assay may be due to metabolic degradation during the longer assay period. Little inhibition of Pr55^{gag} processing in infected cells by compound 4 or pepstatin A was observed (Fig. 2, 'Comp 4' and 'Pep' in lanes 0, 1.5 and 3). The inefficacy of compound 4 is curious as its inhibition constant is comparable to that of compound 1. These results indicate that structural features other than those that impart inhibitory activity to these compounds *in vitro* are required for the blocking of polypeptide processing during maturation of virions in infected cell cultures.

To determine whether or not the virus could establish an acute infection in T lymphocytes in the presence of a protease inhibitor, infectivity studies were conducted in which HIV-1, isolated from a culture of infected H9 cells²⁸ (H9 III_B), was

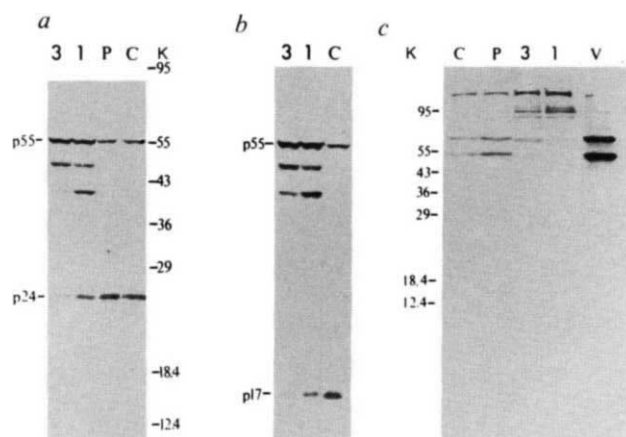


FIG. 1 Effects of HIV-1 protease inhibitors on processing of Pr55^{gag} and Pr160^{gag-pol} proteins in CEM cells chronically infected with HIV-1, strain III_B. Infected cell cultures were treated with either 20 μ M of each inhibitor or with 0.2% dimethylsulphoxide (DMSO) (control) at 8 h intervals for 48 h. Proteins in cell extracts were separated on SDS-polyacrylamide (10% polyacrylamide) gels²⁹ and then electroblotted onto nitrocellulose sheets. Samples were from cells treated with compound 3 (lane 3), compound 1 (lane 1), pepstatin A (lane P), and a 1% DMSO control (lane C). For comparison, a control lane of purified HIV-1 (strain III_B) grown in H9 cells is shown in lane V. Western blots of proteins were probed with monoclonal antibodies (1:1,000) recognizing the Pr55^{gag}/p24 antigen (Beckman AS6) (a), the Pr55^{gag}/p17 antigen (Beckman AS3) (b), and the p66 and p51 subunits of reverse transcriptase protein (Beckman AS6) (c), and then reacted with alkaline phosphatase-labelled goat anti-mouse IgG (1:1,200). Standard relative molecular mass markers for proteins (K) were from Diversified Biotech.

added to uninfected Molt 4 cells immediately before the introduction of inhibitor in a single aliquot. As the individual cultures were diluted sixfold over the course of the experiment by the addition of fresh medium at days 2 and 5, the concentrations of the inhibitors added at day 1 represent an upper limit to their effective concentrations. After 7 days, the extent of infection of the Molt 4 cells was determined by quantification of particle-associated reverse transcriptase activity and p24 antigen in the culture supernatants and by estimating the number of apparent syncytia formed.

As shown in Table 1, the effects of compounds 1-3 on viral infectivity correlated with their ability to inhibit HIV-1 protease in a chronically infected cell culture. Compounds 1 and 3 almost completely inhibited viral infectivity at initial concentrations (25-100 μ M) that also profoundly inhibited the protease in chronically infected cell cultures. Moreover, this diminution of both the reverse transcriptase activity and p24 antigen paralleled a decrease in the number of syncytia formed by these cells during the course of the infection. Compound 2, which may be metabolically unstable in cell culture, exerted little or no effect

on viral infectivity at any concentration. Because one would not expect the inhibition of the viral protease by compounds 1 and 3 to affect the expression and processing of the cell-surface glycoprotein gp120, the observation of diminished syncytia, commensurate with reductions in reverse transcriptase activity and p24 antigen, indicates a substantial suppression of the population of replication-competent virions budding from the lymphocytes that are capable of establishing an acute infection in these cells.

We have shown here that novel peptide analogue inhibitors of HIV-1 protease significantly inhibit the proteolytic processing of Pr55^{gag} and Pr160^{gag-pol} polyproteins in chronically infected T lymphocyte cultures actively producing mature virions, and prevent the ability of replication-competent virus to infect virus-naïve T lymphocytes. These data strongly indicate that the inhibition of HIV-1 protease is responsible for rendering the virus non-infectious. Compounds 1-3 inhibit the aspartic proteases bovine spleen cathepsin D and human renin relatively poorly; the inhibition constant K_i (cathepsin D) and 50% inhibitory concentration IC_{50} (renin) values are, respectively, >3 μ M

TABLE 1 Effects of protease inhibitors on HIV-1 infectivity of T-lymphocytes

Culture components	Additive concentration at Day 1/Day 7	Reverse transcriptase activity (average c.p.m.)	p24 antigen (average ng ml ⁻¹)	Syncytial formation	% Inhibition of infectivity	
					Reverse transcriptase	p24
Molt 4 cells	—	1,500	0.4	—	—	—
Molt 4/virus/DMSO	0.1%/0.017%	128,500	>400	+++++	0	0
	0.25%/0.04%	78,200	>400	+++++	0	0
	1.0%/0.17%	70,400	>400	+++++	0	0
Molt 4/virus/compound 1	10/1.7 μ M	68,000	>400	+++++	47	0
	25/4.2 μ M	9,700	116	++	88	>70
	100/17 μ M	2,900	1.8	—	96	>99
	10/1.7 μ M	93,100	>400	+++++	28	0
Molt 4/virus/compound 2	25/4.2 μ M	72,000	>400	+++++	8	0
	100/17 μ M	71,300	>400	+++++	0	0
	10/1.7 μ M	30,900	>392	+++++	76	>2
Molt 4/virus/compound 3	25/4.2 μ M	1,800	5.4	—	98	>99
	100/17 μ M	1,900	6.8	—	97	>98

Cultures of uninfected Molt 4 (4×10^4 cells in 0.1 ml of RPMI 1640 medium/20% fetal calf serum) were treated with infectious HIV-1, strain III_B (200 infectious units in 0.05 ml), which was recovered in the supernatant of a culture of H9III_B cells²⁸ by filtration through a 0.22 μ m filter. Solutions of compounds 1-3 (10 mM in DMSO) were diluted into medium, and added in a single dose to the infected cell cultures in 0.05 ml aliquots to yield concentrations at day 1 as shown with respective concentrations of 0.1%, 0.25% and 1.0% DMSO. Each treatment was carried out in duplicate and results were averaged. On days 2 and 5 post infection, about 0.5 ml of fresh medium were added to each culture, and on day 7, the reverse transcriptase activity²⁸ and amount of p24 antigen (Dupont HIV-1 Antigen Capture Kit) in each culture were determined. Viral infectivity in the absence of the inhibitors was determined in infected cell cultures treated with 0.1-1% DMSO, and a control sample contained no virus. Syncytial formation (multinucleated cells of >5-times the diameter of a Molt 4 cell) in each well was scored by microscopic examination ($\times 40$), and scoring was based on a graded level of no apparent syncytia (—) to >50 syncytia (5+) in a single microscopic field.

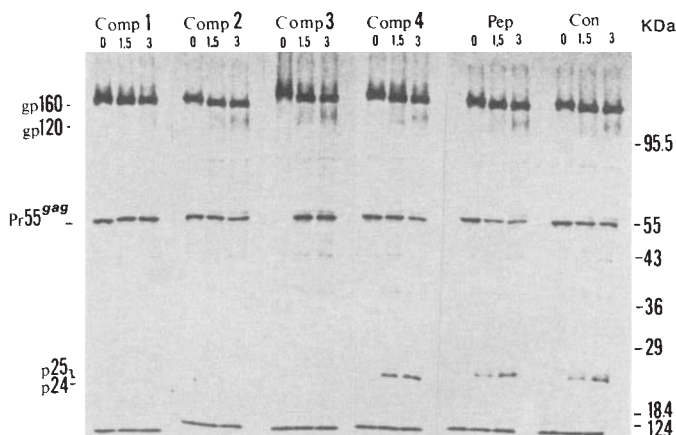


FIG. 2 Pulse-chase analysis of HIV-1/III_b proteins in the presence of protease inhibitors. Individual cultures of chronically infected CEM/III_b cells were treated with 100 μ M each of compounds 1–4 (Comp 1–4), or pepstatin A (Pep) (with control sample (Con) containing 1% DMSO) for 1 h in serum-free, methionine-free Minimum Essential Medium (MEM) before a 20 min pulse-labelling period with [35 S]methionine (100 μ Ci ml $^{-1}$, 1,100 Ci mmol $^{-1}$). Labelled cultures were washed twice with ice-cold PBS and then chased with MEM containing 100 μ M inhibitor, a 100-fold excess of methionine (1.5 mg ml $^{-1}$), and 10% fetal bovine serum for 0, 1.5 and 3 h. Cells were collected at each time point by centrifugation and then lysed in 0.02 M Tris-HCl buffer, pH 7.5, 0.15 M NaCl, 0.001 M EDTA, 0.5% Triton X-100, 0.1% SDS. [35 S]methionine-labelled HIV-specific proteins were immunoprecipitated using pooled AIDS patient sera. Antigen-antibody complexes were recovered with IgG-Sorb (The Enzyme Center, Boston) as previously described³⁰. Proteins were resolved by electrophoresis on SDS-polyacrylamide gels (10% polyacrylamide) and visualized by autoradiography. Immunoprecipitated HIV-1 proteins in control (Con) lanes were *env* gp160, Pr55^{gag} and p24. Protein relative molecular mass (K) markers are shown.

and >100 μ M. From the standpoint of potential therapeutic applications, this differential inhibition of the viral and mammalian aspartic proteases is a desirable property in that such compounds might demonstrate low toxicity *in vivo*. Indeed, no overt cytotoxicity, as determined by trypan blue staining of uninfected CEM cells, was observed for compounds 1–4 at concentrations as high as 100 μ M over a 48-h period (data not shown). Studies are underway to assess the effect of these inhibitors on virion morphology. □

28. Popovic, M., Sarngadharan, M., Read, E. & Gallo, R. *Science* **224**, 497–500 (1984).
29. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
30. Lambert, D. M. & Pons, M. W. *Virology* **130**, 204–214 (1983).

ACKNOWLEDGEMENTS. We thank Dr Martin Rosenberg for his support of this work, Drs James Young and Jeffrey Leary for discussions, Ken Strahs for the gift of monoclonal antibodies, and Dr Kei Lei Fong for providing pooled AIDS patient sera.

Structure and evolution of a human erythroid transcription factor

Cecelia D. Trainor, Todd Evans, Gary Felsenfeld & Mark S. Boguski*

Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892, USA

* National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894, USA

VERTEBRATE erythroid cells contain a tissue-specific transcription factor referred to as Eryf 1 (ref. 1), GF-1 (ref. 2) or NF-E1 (ref. 3), for which binding sites are widely distributed in the promoters and enhancers of the globin gene family^{1–8}, and of other erythroid-specific genes^{9,10}. Aberrant binding of the human factor to a mutant site has been implicated in one form of hereditary persistence of fetal haemoglobin (HPFH; ref. 2). The complementary DNAs for both the chicken cEryf 1 (ref. 11) and mouse mEryf 1 (ref. 12) encoding genes have recently been cloned. We report here the cloning of the cDNA for the human Eryf 1 encoding gene. The central third of the hEryf 1 cDNA, containing two 'finger' motifs, is almost identical to that of chicken or mouse. The amino- and carboxy-terminal thirds of the human protein are similar to those of mouse, but are strikingly different from the corresponding domains in chicken. The evidence indicates that these erythroid regulatory factors evolved from a common precursor composed of two distinct kinds of repeated domains, which subsequently evolved at greatly different rates.

Several independent hEryf 1 clones were isolated by screening a human bone marrow cDNA library with a cDNA clone of the chicken Eryf 1 encoding gene (ref. 11). The largest insert obtained was 1,513 nucleotides (nt), which by northern analysis (Fig. 1a) hybridizes to a predominant 1,550-nt message in K562 total-cell RNA, and to a 1 kilobase (kb) message in chicken erythrocyte RNA; this message was not observed in RNA isolated from nonerythroid cells, consistent with results previously reported using as a probe the chicken or mouse cDNA^{11,12}.

The human and chicken cDNA clones were transcribed *in vitro* and the RNAs translated in a rabbit reticulocyte lysate in the presence of [35 S]methionine. The protein generated from the human clone migrates with a relative molecular mass (M_r) of ~49,000 (49K) when analysed by SDS-PAGE (data not shown) consistent with the size reported for the human protein purified from K562 cells (50K; ref. 12). The *in vitro* translation products were used in a gel mobility-shift assay (Fig. 1b). A specific complex is formed when lysate primed with RNA from the hEryf 1 clone (H) is added to DNA oligomers containing an Eryf 1 binding site derived from either the chicken (lane 1) or human (lane 2) β -globin enhancers. The complex is identical in mobility to a specific complex formed in the presence of K562 nuclear extract (lanes 5–12). The translation product from the cEryf 1 clone shows identical binding specificity (C).

Southern blots of chicken or human genomic DNA were probed with the corresponding cDNA clones (Fig. 1c), which hybridize predominantly to a single restriction fragment, or to two smaller fragments. In addition, the human cDNA clone hybridizes most strongly to the same chicken genomic restriction fragments as does the chicken cDNA clone (arrows), which indicates that the two cDNAs represent the same, most probably

1. Gallo, R. C. & Montagnier, L. *Scient. Am.* **259**, 40–48 (1988).
2. Ratner, L. et al. *Nature* **313**, 277–284 (1985).
3. Sanchez-Pescador, R. et al. *Science* **227**, 484–492 (1985).
4. Wain-Hobson, S. et al. *Cell* **40**, 9–17 (1985).
5. Muesing, M. A. et al. *Nature* **313**, 450–458 (1985).
6. Jacks, T. et al. *Nature* **331**, 280–283 (1988).
7. Veronese, F. D. et al. *AIDS Res. human Retrovir.* **3**, 253–262 (1987).
8. Veronese, F. D. et al. *J. Virol.* **62**, 795–801 (1988).
9. Henderson, L. E., Copeland, T. D., Sowder, R. C., Schultz, A. M. & Oroszian, S. in *Human Retroviruses, Cancer, and AIDS: Approaches to Prevention and Therapy* 135–147 (Alan R. Liss, New York, 1987).
10. Kohl, N. E. et al. *Proc. natn. Acad. Sci., U.S.A.* **85**, 4686–4690 (1988).
11. Peng, C., Ho, B., Chang, T. & Chang, N. *J. Virol.* **63**, 2550–2556 (1989).
12. Gottlinger, H., Sodroski, J. & Haseltine, W. *Proc. natn. Acad. Sci., U.S.A.* **86**, 5781–5785 (1989).
13. Seelmeier, S., Schmidt, H., Turk, V. & von der Helm, K. *Proc. natn. Acad. Sci., U.S.A.* **85**, 6612–6616 (1988).
14. Crawford, S. & Goff, S. *J. Virol.* **53**, 899–907 (1985).
15. Katoh, I. et al. *Virology* **145**, 280–292 (1985).
16. Debouck, C. et al. *Proc. natn. Acad. Sci., U.S.A.* **84**, 8903–8906 (1987).
17. Meek, T. D. et al. *Proc. natn. Acad. Sci., U.S.A.* **86**, 1841–1845 (1989).
18. Pearl, L. H. & Taylor, W. R. *Nature* **329**, 351–354 (1987).
19. Navia, M. A. et al. *Nature* **337**, 615–620 (1989).
20. Wlodawer, A. et al. *Science* **245**, 616–621 (1989).
21. Holladay, M. W., Salituro, F. G. & Rich, D. H. *J. med. Chem.* **30**, 374–383 (1987).
22. Szekle, M., Jones, D., Atrash, B., Hallett, A. & Leckie, B. in *Peptides Structure and Function. Proceedings of Eighth American Peptide Symposium* (eds Hruby, V. J. & Rich, D. H.) 579–582 (Pierce Chemical Co., Rockford, Illinois, 1983).
23. Dreyer, G. B. et al. *Proc. natn. Acad. Sci., U.S.A.* (in the press).
24. Moore, M. L. et al. *Biochem. Biophys. Res. Comm.* **159**, 420–425 (1989).
25. Darke, P. L. et al. *Biochem. Biophys. Res. Comm.* **156**, 297–303 (1988).
26. Matthews, T. J. et al. *Proc. Acad. natn. Sci., U.S.A.* **84**, 5424–5428 (1987).
27. Mervis, R. J. et al. *J. Virol.* **62**, 3993–4002 (1988).

single-copy, gene. Additional fainter bands are observed in the genomic DNA of both species and could represent other, more distantly related, genes.

The nucleotide sequence and deduced amino-acid sequence of the hEryf 1 cDNA is presented in Fig. 2. It contains a single open reading frame (M_r 42,995) of sufficient size to encode the observed *in vitro* translation product. The amino-acid sequence of hEryf 1 is 86% identical to the mouse protein. The central part of the molecule contains a repeated 'finger' motif of the form Cys-X-Asn-Cys-X₄-Thr-X-Leu-Trp-Arg-Arg-X₃-Gly-X₃-Cys-Asn-Ala-Cys that is also found twice in the mouse and chicken proteins.

We have used a quantitative comparison matrix analysis¹³ to map the repeated sequences within these proteins (Fig. 3a). The location of the tandem finger domains is seen in the self-comparison matrices of both the chicken and human sequences. The human protein has a longer N-terminal region preceding the finger domains and lacks the distal zone of oligoglycine and oligoproline repeats present in the chicken sequence between residues 225 and 299. Both the chicken and the human Eryf 1 sequences contain a second system of long internal repeats, among the first 50 and last 50 residues of the proteins. In the human sequence, however, there appears to be a third copy centered around residues 135–180. Detailed analysis of this second system of repeats (Fig. 3b) reveals that in addition to the tandem finger domains, the human protein contains three divergent copies of an approximately 94-residue sequence. The modular structure of hEryf 1 is summarized in Fig. 3c.

A comparison matrix for chicken versus human Eryf 1 (Fig. 3a) illustrates that these sequences are clearly related colinearly in their respective tandem finger domains, as well as less markedly in their C-terminal regions. Traces of the second repeat system are even evident in this intersequence comparison. But a multiple alignment using all three proteins (Fig. 4) indicates

TABLE 1 Rates of nucleotide substitution in the erythroid transcription factor gene: comparison with other genes

Coding sequence	Synonymous rate	Nonsynonymous rate
Erythroid factor:		
complete sequence	4.15 (0.44)	0.449 (0.06)
N-terminal domain	3.35 (0.54)	0.673 (0.10)
finger domain	4.85 (0.95)	0.068 (0.04)
C-terminal domain	5.42 (0.13)	0.537 (0.13)
Histone H4	6.13 (1.32)	0.027 (0.03)
Histone H2B	3.59 (0.69)	0.076 (0.04)
Haemoglobin alpha	3.94 (0.60)	0.56 (0.09)
Haemoglobin beta	2.96 (0.46)	0.87 (0.11)
Mammalian average	4.65 (2.06)	0.88 (0.75)

The codon alignments used to compute the rates are based on the amino-acid sequence alignments in Fig. 4. Rates are expressed as the number of substitutions per site per 10^9 years (standard errors are in parentheses) and different mammalian orders are assumed to have diverged 80 million years ago. The rates for the Eryf 1 gene sequence were computed according to Li *et al.*¹⁴; all other rates were taken from Table 2 (ref 14). The N-terminal, finger and C-terminal domains in the human and mouse proteins correspond to nucleotides 1–588, 589–951 and 952–1239 in the coding sequences of their respective mRNAs. The nonsynonymous rate of substitution for chicken and human finger domains was 0.129×10^{-9} (with a standard error of 0.003×10^{-9}) based on a divergence time of 270 million years ago. Meaningful rates for the N- and C-terminal regions could not be determined owing to their much greater divergence¹⁴.

that there are several large deletions in the N-terminal region of the chicken sequence, as well as additional minor deletions in the C-terminal domain.

We have analysed the nucleotide substitution rates in the human and mouse messenger RNA sequences (Table 1). The synonymous (phenotypically silent) rate of nucleotide substitu-

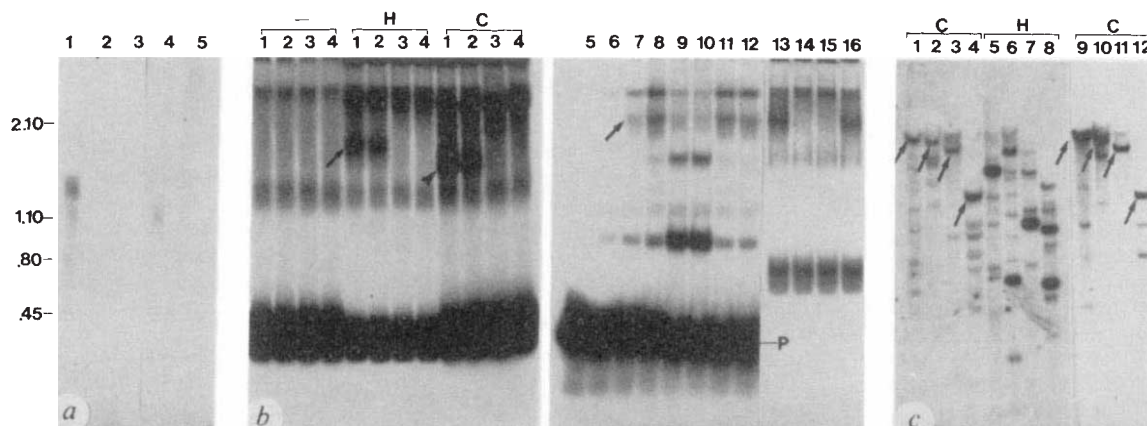


FIG. 1 Characterization of human Eryf 1 cDNA. a, Northern analysis using a human Eryf 1 cDNA clone as a probe. Blots contained: lane 1, 40 μ g K562 total-cell RNA; lane 2, 40 μ g H9 (a human T-cell line) total-cell RNA; lane 3, 2 μ g poly(A)-selected RNA from human brain. Lanes 4 and 5 contained 30 μ g total-cell RNA isolated from 9-day embryonic chicken erythrocytes or brain tissue, respectively. β -actin mRNA was detected in all lanes (data not shown). b, Gel mobility-shift assays using *in vitro* translation products of the Eryf 1 cDNA clones. Labelled ($[^{35}\text{S}]$ methionine) translation products and unlabelled oligonucleotide probes were used (lanes 1–4 and 13–16) to avoid detection of endogenous Eryf 1 in the lysate. Probes: Ch1 (lane 1) contains a strong Eryf 1 binding site derived from the chicken β -globin enhancer¹¹; Hm3 (lane 2) and Hm2 (lane 3) contain sequences from the human β -globin enhancer; only Hm3 binds strongly to Eryf 1 (ref 1). The Ch2 probe (lane 4) does not contain an Eryf 1 binding site (unpublished data). Probes (0.4 pmol) were incubated with unprimed lysate (–), or lysate primed by either hEryf 1 RNA (H) or cEryf 1 RNA (C). The cEryf 1-DNA complex is of faster mobility, consistent with the smaller size of the chicken protein (38K, refs 1, 11). To compare the *in vitro* translation product with hEryf 1 (lanes 5–8), ^{32}P end-labelled Ch1 probe (P) was incubated with (0, 0.5, 1.0

or 2 μ l, respectively) of K562 cell nuclear extract before electrophoresis. Lanes 9 and 10 are as in lane 8 except that a 20-fold or 40-fold excess of Ch1 oligomer was included. In lanes 11 and 12, similar amounts of the nonspecific Ch2 oligomer were used. Lanes 13–16 were run on the same gel as lanes 5–12 for direct comparison: labelled hEryf 1 translation product was incubated with unlabelled Ch1, Hm2, Ch2 and Hm3, respectively. Arrows indicate the hEryf 1-DNA complex; the arrowhead shows the cEryf 1-DNA complex. c, Southern analysis. Chicken (C) or human (H) genomic DNA (10 μ g per lane) was digested with *Sac*I, *Bam*HI, *Eco*RI or *Pst*I (left to right); blots were probed with labelled hEryf 1 (lanes 1–8) or cEryf 1 DNA (lanes 9–12). Arrows indicate the predominant restriction fragments which are detected by both probes in chicken DNA.

METHODS. A human bone marrow cDNA library in λ gt11 was constructed from mRNA derived from a single adult male (Clontech), and hybridized to a cEryf 1 cDNA probe in $2 \times \text{SSC}$, 0.1% SDS at 55 $^{\circ}\text{C}$. For northern blots RNA was electrophoresed on formaldehyde gels; blots were hybridized and washed¹⁹ at 65 $^{\circ}\text{C}$ (lanes 1–3) or 55 $^{\circ}\text{C}$ (lanes 4–5). DNA blots were hybridized at 42 $^{\circ}\text{C}$ in 50% formamide, 10% dextran sulphate, 1% SDS and 1 M NaCl, and washed in $2 \times \text{SSC}$, 1% SDS at 60 $^{\circ}\text{C}$.

Given the known properties of regulatory proteins, it seems likely that a large part of the DNA-binding specificity of Eryf 1 resides in finger domains F1 and F2 (Fig. 3c, ref. 11). The Cys-X-X-Cys sequence present at both ends of each Eryf 1 finger is also found in fingers of the steroid-receptor-superfamily DNA-

We speculate that other domains of Eryf1 might interact directly with components of the transcriptional apparatus. The second family of internal repeats (R 1-3 in human and mouse) may serve this purpose. The corresponding repeat system in the chicken protein is disrupted by multiple deletions; the relaxed

We speculate that other domains of Eryf1 might interact directly with components of the transcriptional apparatus. The second family of internal repeats (R 1-3 in human and mouse) may serve this purpose. The corresponding repeat system in the chicken protein is disrupted by multiple deletions; the relaxed

FIG. 2 The nucleotide and deduced amino-acid sequence of the human Eryf 1 cDNA. The nucleotide sequence of both strands of the largest Eryf 1 clone was determined by the chain-termination method using T7 DNA polymerase (Sequenase USB). The entire sequence of an independent 1,500-nt clone and partial sequences of two smaller clones were also determined and are in complete agreement. The cysteine residues that are components of the tandem finger domains are bracketed and a putative polyadenylation signal is underlined.

FIG. 3 Amino-acid sequence analyses. *a*, Protein comparison matrices. Program CMPSEQ84 (ref. 13) was used to compute the comparison matrices with a window of 53 residues using the PAM250 matrix²⁰ for scoring. A plotting threshold corresponding to a chance occurrence of $\leq 10^{-3}$ was used throughout; the highest-scoring diagonals (corresponding to the tandem finger domains) achieved scores with chance probabilities of the order of 10^{-20} . A chance probability of 0.53×10^{-10} was obtained by this method for the most significant internal repeats in *Xenopus* transcription factor IIIA (ref. 21). Grid marks correspond to 50 amino-acid intervals; numbering is relative to the initiator methionine of either cEryf 1 (1–304) or hEryf 1 (1–413). The significance of the internal repeats in the chicken, human and mouse proteins was confirmed using intrasequence RELATE analysis²⁰; the scores (in s.d. units) for the self-comparisons were 15.0, 10.9 and 9.8, respectively (scores of ≥ 3.0 are diagnostic of internal duplication). *b*, Human erythroid factor: alignments of internal repeats. A new dynamic programming method²² was used to optimally align the two sets of internal repeats in the human sequence. To analyse the second repeat system, the N-terminal region (residues 1–195) was split into two sub-sequences which were simultaneously aligned with the C-terminal region (residues 320–413). sequences which were simultaneously aligned with the C-terminal region (residues 320–413). Schematic diagram of repeated sequence organization based on the analyses in panels *a* and *b*. The symbols C..C refer to the locations of cysteine subsequences in the 'finger' domains. The boxed plus signs (+) indicate positive charge clusters (as defined in reference 23) corresponding to residues 232–253 and 287–317 in the human and mouse sequences. The chicken sequence also has corresponding positive charge clusters associated with the C-terminal regions of its 'finger' repeats (data not shown).

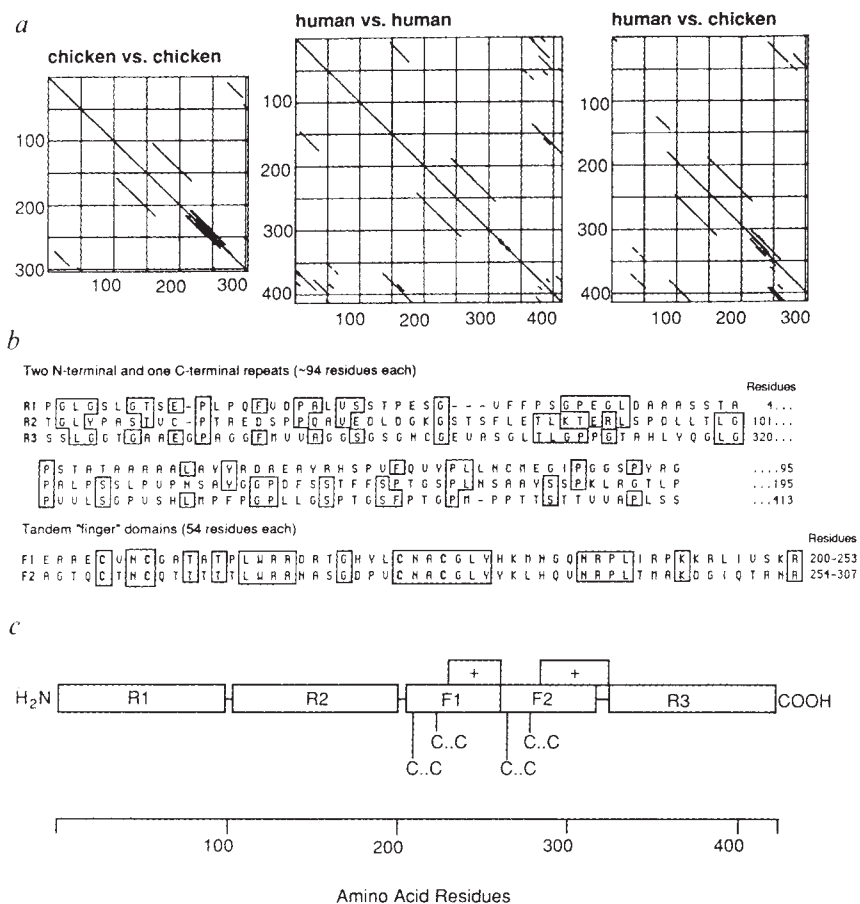


FIG. 4 All three sequences were simultaneously aligned²² using appropriate weights²⁴ to control for the greater pairwise similarity of the human and mouse sequences. Tandem finger domains are underlined. Numbering corresponds only to the human and mouse proteins. Single-letter amino-acid code is used.

```
MEFVALGGPDAGSPTP-FPD-----EAGAF-----GLGGGERTEA----- chicken
MEFPGLGSLGTSEPLPQFVDPALVSSPTESGVFPSPGPEGLDAAASSTAPSTATAAAAAAL human
MDFPLGLGALGTSEPLPQFVDSALVSSPDSTGFFSSGPEGLDAAASSTSPNAATAAASAL mouse
    10      20      30      40      50      60

-----GGLIAYPPSGRVSLVP-----WA-DTGTGTGTPQWVPPATQMEPPH-
AAYYDAEAYRHSPVFQVYPLNCMEGIPGGSPYAGWAYGKTGLYPASTVCPTRDSSPPQA
AAYYREAEAYRHSPVFQVYPLNSMEGIPGGSPYASWAYGKTALYPASTVCPSHEDAPSQA
    70      80      90      100     110     120

-----YLELLQPPRGSPHPSSGPLL-----PL
VEDLDGKGSTSFLETLKTERLSPDLLTLGPALPSSLPVNSAYGGPDFSTFFSPTGSP
LEDQEGKSNNTFLDTLKTTERLSPDLLTLGTALPASLPVTCGAYGGADFSPFFSPTGSP
    130     140     150     160     170     180

SSGP-----PPCEARECVNCGATATPLWRRDGTGHYLCNACGLYHRLNGONRP
NSAAYSSPKLRGTLP LPPCEARECVNCGATATPLWRRDGTGHYLCNACGLYHKMNGONRP
SSAAYSSPKFHGSLPLAPCEARECVNCGATATPLWRRDGTGHYLCNACGLYHKMNGONRP
    190     200     210     220     230     240

LIRPKKRLIVSKRAGTVCNCOTSTTTTLWRRSPMGDPVCNACGLYKLVHVNRLTMRKD
LIRPKKRLIVSKRAGTVCNCOTSTTTTLWRRNAGDPVCNACGLYKLVHVNRLTMRKD
LIRPKKRMIVSKRAGTVCNCOTSTTTTLWRRNAGDPVCNACGLYKLVHVNRLTMRKD
    250     260     270     280     290     300

GIOTRNRKVSSKGGKRRPP---GGGNPSATAGGGAPMGGGDPSPM-PPPPPPPPAAAPPQS
GIOTRNRKASGKGKKRGSSSLGGTGAAGPAGGFVVAGSGSGNCGEASGLTLGPPGT
GIOTRNRKASGKGKKRGSSNLAGAGAAEGPAGGFVVAGSSSGNCGEASGLALGTAGT
    310     320     330     340     350     360

DALY-ALGPVVLGSG---HFLPF-----GNSGGFFGGGAGGYTAPPGLSPQI--
AHLYQGLGPVVLGSPVSHLMPFPGPLLGSPPTGSPPTGMPPTTS TT VVAPLSS
AHLYQGLGPVVLGSPVSHLMPFPGPLLGSPPTGSPPTGMPPTTS TT VVAPLSS
    370     380     390     400     410
```

evolutionary constraints which this implies (demonstrated also in Table 1) may reflect changes in the nature of the other components with which they could interact. Alternatively, there may be less stringent requirements for conservation of these domains, compared to the DNA-binding domain. It may be relevant that the human cDNA, though active, is less efficient than the chicken clone in a transient transactivation assay carried out in nonerythroid chicken cells (unpublished data). □

Received 11 September; accepted 13 November 1989.

1. Evans, T., Reitman, M. & Felsenfeld, G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5976-5980 (1988).
2. Martin, D. I. K., Tsai, S. & Orkin, S. H. *Nature* **338**, 435-438 (1989).
3. Wall, L., deBoer, E. & Grosfeld, F. *Genes Dev.* **2**, 1089-1100 (1988).
4. de Boer, E., Antoniou, M., Mignotte, V., Wall, L. & Grosfeld, F. *EMBO J.* **7**, 4203-4212 (1988).
5. Superti-Furga, G., Barberis, A., Schaffner, G. & Busslinger, M. *EMBO J.* **7**, 3099-3107 (1988).
6. Galson, D. L. & Housman, D. E. *Molec. cell. Biol.* **8**, 381-392 (1988).
7. Barnhart, K. M., Kim, C. G. & Sheffery, M. *Molec. cell. Biol.* **9**, 2606-2614 (1989).
8. Perkins, N. D., Nicolas, R. H., Plumb, M. A. & Goodwin, G. H. *Nucleic Acids Res.* **17**, 1299-1314 (1989).
9. Mignotte, V., Wall, L., de Boer, E., Grosfeld, F. & Romeo, P.-H. *Nucleic Acids Res.* **17**, 37-54 (1989).
10. Plumb, M. et al. *Nucleic Acids Res.* **17**, 73-91 (1989).
11. Evans, T. & Felsenfeld, G. *Cell* **58**, 877-885 (1989).
12. Tsai, S. et al. *Nature* **339**, 446-451 (1989).
13. McLachlan, A. D. *J. molec. Biol.* **169**, 15-30 (1983).
14. Li, W.-H., Wu, C.-I. & Luo, C.-C. *Molec. Biol. Evol.* **2**, 150-174 (1985).
15. Evans, R. M. *Science* **240**, 889-895 (1988).
16. Beato, M. *Cell* **56**, 335-344 (1989).
17. Kumar, V. & Chambon, P. *Cell* **55**, 145-156 (1988).
18. Arst, H. N. Jr, Kudla, B. & Martinez-Rossi, N. *Trends Genet.* **5**, 291 (1989).
19. Church, G. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991-1995 (1984).
20. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (National Biomedical Research Foundation, Washington, D.C., 1978).
21. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609-1614 (1985).
22. Lipman, D. J., Altschul, S. F. & Kececioglu, J. D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4412-4415 (1989).
23. Brendel, V. & Karlin, S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5698-5702 (1989).
24. Altschul, S. F., Carroll, R. J. & Lipman, D. J. *J. molec. Biol.* **207**, 647-653 (1989).

ACKNOWLEDGEMENTS. We thank W.-H. Li, V. Brendel and S. Karlin for providing computer programs, and A. Dean and Q.-H. Gong for K562 nuclear extract. The sequence data in this paper has been submitted to the (EMBL/GenBank/DBJ) database under accession number X17254.

GUIDE TO AUTHORS

Please follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers for whom English is a second language and those who work in other fields. Please write clearly and simply, avoiding unnecessary technical terminology. *Nature's* staff will edit manuscripts to those ends if necessary. Contributors should check their proofs carefully. Because of the competition for space, many of the papers submitted for publication cannot be accepted. For this reason, and because brevity is a great assistance to readers, papers should be as brief as is consistent with intelligibility. Please note that one printed page of *Nature*, without diagrams or other interruptions of the text, has fewer than 1,300 words.

CATEGORIES OF PAPER

Review Articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance but must not exceed six pages of *Nature*. **Articles** are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles should be accompanied by a heading of 50-80 words written to advertise their contents in general terms, to which editors will pay particular attention. A heading (printed in italic type) is not an abstract and should not usually contain numbers or measurements. The study should be introduced in more detail in the first two or three paragraphs, which should also briefly summarize its results and implications.

Articles may contain a few subheadings of two or three words. The meaning of the text should not depend on the subheadings, whose function is to break up the text and to point to what follows. There should be fewer than 50 references.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should not have more than 1,000 words of text or more than four display items and should not occupy more than two pages of *Nature*. The first paragraph should describe, in not more than 150 words, the origins and chief conclusions of the study. Letters should not have subheadings or more than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but some are unsolicited. They are normally between one and four pages of *Nature* in length.

News and Views articles are intended to inform non-specialist readers about a recently published advance. Suggestions should be made to the News and Views Editor. Illustrations are welcome. Proposals for meeting reports should be agreed in advance.

Scientific Correspondence is for the discussion of scientific matters, including contributions published in *Nature*. Priority is given to contributions of less than 500 words and five references. Figures are welcome.

Submission. All manuscripts may be submitted to either London or Washington. They should not be addressed to editors by name. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax (VAT).

GENERAL

All contributions submitted for publication in *Nature* should conform with these rules.

Manuscripts should be typed, double-spaced, with a good-quality printer, on one side of the paper only. Four copies are required, each accompanied by lettered artwork. Four copies of half-tones should be included. Reference lists, figure legends and so on should be on separate sheets, all of which should be double-spaced and numbered. Copies of relevant manuscripts in press or submitted for publication elsewhere should be included, clearly marked as such. Revised and resubmitted manuscripts should also be clearly marked as such and labelled with their reference numbers.

Titles should say what the paper is about with the minimum of technical terminology and in fewer than 80 characters. Authors should avoid active verbs, numerical values, abbreviations and punctuation and should include one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name, figure number and, when known, manuscript reference number. Original artwork should be unlettered. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are permitted only if the parts are closely related, either experimentally or logically. Suggestions for cover illustrations, with captions, are welcome. Original artwork (and one copy of the manuscript) will be returned when a manuscript cannot be published.

Colour Artwork. A charge of £500 a page is made for 4-colour figures. Inability to meet these costs will not prevent the publication of essential colour figures if the circumstances are explained.

Figure legends must not exceed 300 words and ideally should be much shorter, and should use telegraphic form wherever possible. The figure should be described first, then, briefly, the method. Reference to a method described elsewhere is preferable to a full description. Methods should not be described in detail in the text.

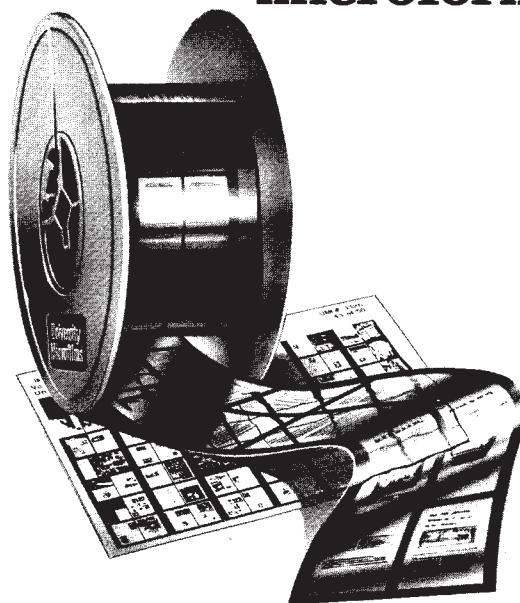
References should be numbered sequentially as they appear in the text, followed by those in tables and finally those in the figure legends. Reference numbers apply only to papers published or in the press, and a different number should be given to each paper cited. All other forms of reference (including unrefereed abstracts) should be included in the text as personal communication, manuscript in preparation or preprint (with number and institution where appropriate). Text should not be included in the reference list. References should be abbreviated according to the *World List of Scientific Periodicals* (fourth edition, Butterworth, London, 1963-65). The first and last page numbers should be cited. Reference to books should clearly indicate the publisher and the date and place of publication.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided as much as possible and, when used, defined. Editors will shorten words if necessary. In case of doubt, SI units should be used.

Footnotes should not be used except for changed addresses or to identify the corresponding author if different from the first-named.

Acknowledgements should be brief. Grant numbers and contribution numbers are not allowed.

nature is available in microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

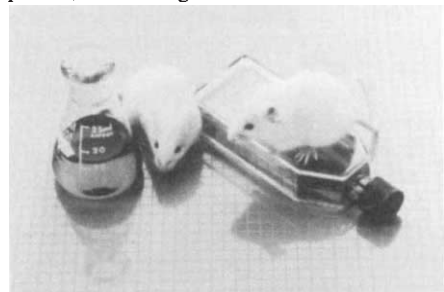
Phone () _____

University Microfilms International

New year product bonanza

Products for the New Year include variable-intensity transilluminators for increased user safety, an *in vivo* cell-tracking kit and a monoclonal antibody for the localization of muscle-specific actin.

TACONIC Farms, Inc. has entered into an agreement with Fox Chase Cancer Centre to breed and distribute **severe combined immune deficient (SCID) mice** (*Reader Service No. 100*). According to Taconic Farms, the SCID mice can be used as models for testing cancer treatments because they readily accept tissue transplants, including human tumours. The

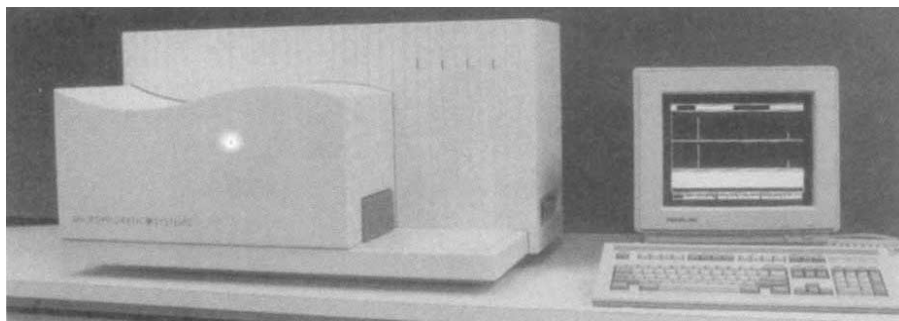


Test cancer treatments with SCID mice from Fox Chase Cancer Centre.

foundation breeder stock from Fox Chase will be hysterectomy-derived under germ-free conditions prior to distribution. Due to the susceptibility of SCID mice to opportunistic infections, colonies are maintained in gnotobiotic isolators. SCID inbred mice and SCID control strain C.B-17/Icr congenic mice will be available in the US and Canada in early 1990. Fox Chase is also licensing two other firms to distribute SCID mice in Europe and Israel.

Variable-intensity transilluminators that limit UV exposure to the user are now available from Fisher Scientific (*Reader Service No. 101*). Three models, each with a different size filter area, are on sale: 8 × 8 inch, 6 × 14 inch and 8 × 16 inch. Each transilluminator has a maximum UV intensity of 8,000 $\mu\text{W}/\text{cm}^2$ and is adjustable from 100% to 20% intensity, allowing detection of bands while minimizing UV damage to the user. The transilluminators operate at 312 nm, which, Fisher says, virtually eliminates DNA degradation caused by 254-nm UV light. All models feature LongLife filter glass, and are fitted with a UV-transmitting filter cover that shields the filter from damage and a hinged UV-blocking safety cover for protection of the user during gel viewing and manipulation.

For an automated high-performance **capillary electrophoresis system** in the mid-price range, Microphoretic Systems recommends the Model 1600 capillary



Model 1600 capillary electrophoresis instrument from Microphoretic Systems.

electrophoresis instrument (*Reader Service No. 102*). Designed for high-sample throughput, the Model 1600 can accommodate a 96-well microtitre plate or eight autosample vials. Up to nine injections can be performed per vial using either electromigration or controlled-vacuum injection methods. The UV photodiode array detection system is designed specifically for capillaries and has five fixed wavelengths with centre bands at 197, 213, 230, 263 and 280 nm. The \$39,500 (US) Model 1600 has a programmable 30-kV power supply, nine analysis methods that can be used with up to four buffers, and programmable buffer/electrolyte replenishment and cycling.

Cell science

Zynaxis Cell Science, Inc. has introduced PKH 26 Cell Linker Kits for *in vivo* cell tracking applications (*Reader Service No. 103*). The red fluorescent molecule (probe) will label various cell types, including red and white cells, platelets, and bone marrow cells. According to

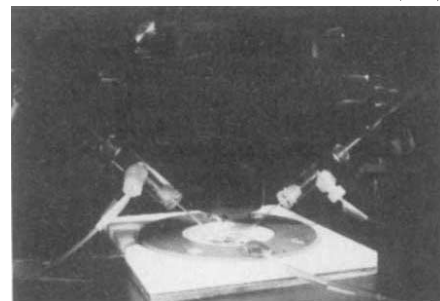


The PKH 26 cell linker, for tracking cells *in vivo*, comes in kit form from Zynaxis.

Zynaxis, the probe binds irreversibly into the lipid region of the cell without affecting the integrity, lifespan or function of the cell. The PKH 26 cell linker has excitation and emission spectra at 488 nm and 570 nm, respectively, and a half life of greater than 100 days in rabbit red blood cells *in vivo*. When used with the PKH 2 green fluorescent compound, the inter-

active cellular behaviour of two cell types can be observed simultaneously. Kits, costing \$295 (US), are available for general cell membrane labelling or for *in vitro* and *in vivo* labelling of phagocytic cells.

Biophysica Technologies, Inc. has introduced a **microscope tissue chamber** for use with an inverted microscope (*Reader Service No. 104*). The \$395 (US)



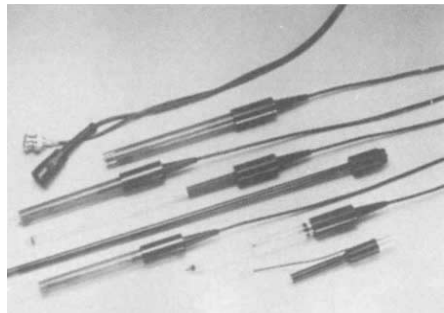
Microscope tissue chambers with thermoregulation accessories are new from Biophysica.

The chamber is designed to allow micro-manipulation, perfusion, atmospheric control and long-term studies under sterile conditions. The ultrathin (0.07 mm) coverslip well bottom allows the use of a short working distance objective, and also fits a standard 35-mm culture dish. Optional thermoregulation accessories are available for between \$185–1,635 (US).

New **feeder cells** for promoting the growth of fastidious cell lines are available from the American Type Culture Collection (*Reader Service No. 105*). Cells from the Swiss albino 3T3 mouse embryo cell line are irradiated so that division is halted but metabolism is maintained. The feeder cells are supplied as \$39 (US) frozen 1-ml aliquots, which, the company says, is sufficient to seed up to 150 cm^2 of surface area.

Measuring devices

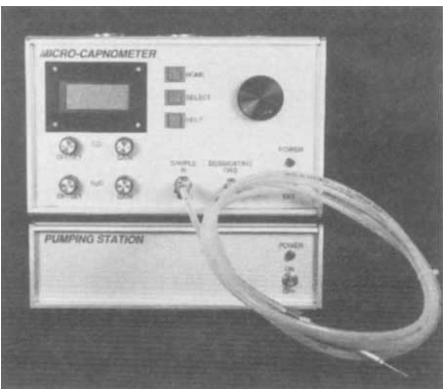
Omega Engineering has introduced a complete line of **pH electrodes** for laboratory use (*Reader Service No. 106*). The PHE-4200 Series gel-filled, combination electrodes are available in five styles: general purpose, double junction, flask size, flat-surface/double-junction and semi-micro. The electrodes measure over



All-purpose pH electrodes from Omega Engineering.

the full pH range at temperatures between 0 and 100 °C and feature a Ag/AgCl reference and polypropylene liquid junction. The electrodes, costing between \$49-84 (US), are supplied in a soaker storage bottle and come with a three-foot cable and BNC connector — although Omega can supply any type of connector.

Low air-sampling rates and fast response times are features of the micro-Capnometer small sample **end-tidal CO₂ and N₂O computer**, says Columbus Instruments Corp. (*Reader Service No. 107*). The microCapnometer, which is designed to measure CO₂ and N₂O percentages and the respiration rate of small animals such



Columbus Instruments' microCapnometer.

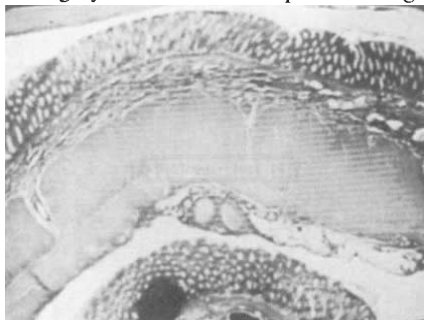
as rats, guinea pigs and rabbits, may have uses in pharmacology and physiology. The instrument requires a sample volume of only 5 ml per minute, which, Columbus says, is considerably less than the 150 ml per minute required by similar instruments. The microCapnometer has a LCD display for showing CO₂ and N₂O waveforms and two linearized outputs for recording waveforms on a strip-chart recorder. The instrument, when purchased with an IBM PC/AT interface, comes complete with software for controlling data acquisition and display.



Eltime Vision Systems' video image archival system: the complete image organiser.

Video view

Capture images from videotape recorders and video cameras for later reference using the **video image archival system (VIAS)** from Eltime Vision Systems (*Reader Service No. 108*). The VIAS can store on optical disk up to 800 colour images with a resolution of 512 × 256 pixels and 32,000 shades of colour. Menu-driven software allows the user to define a coding system for each captured image,



Capture video images with VIAS.

set up a directory of all images, display images on screen with text overlay, and view images with the split-screen facility. The £6,450 (UK) system comprises IBM PC or compatible with 30-MByte hard disk, image IV high-resolution colour frame store, optical disk drive, and 14-inch colour monitor.

Following a hardware and software upgrade, the Videometric-150 colour **image analysis system** from American Innovation, Inc. now features new design functions and applications software (*Reader Service No. 109*). The V-150 can, says the company, distinguish between 16 million true colours in real time, and has applica-

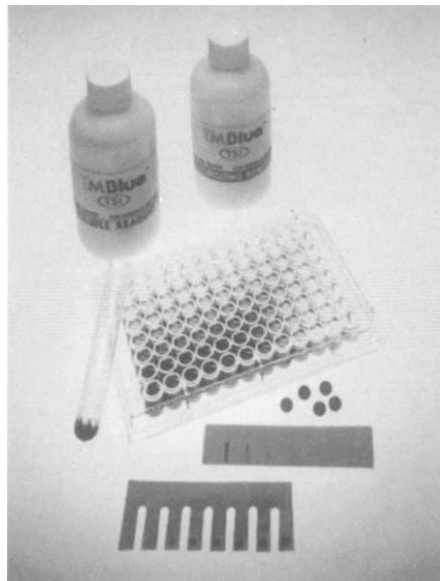


Software for microcirculation studies and autorad analysis — see the V-150.

tions in microscopic, videotape and live image analysis. Edge enhancement algorithms, zoom and pan functions, user-selected colour for tracings, and an expanded macro command language are a selection of the new features. CellTrak and AutoRad applications software can be used for microcirculation studies or autoradiography analysis. Prices for the V-150 start at \$21,500 (US), which includes frame buffer, mouse, 12-inch colour monitor and manual.

Detection systems

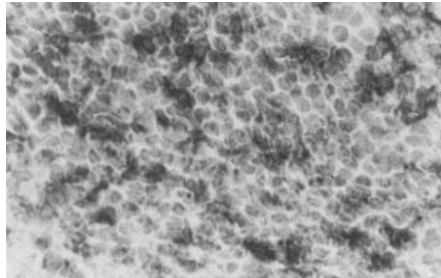
TMBlue is the new diagnostic **substrate for horseradish peroxidase-linked assays** from Transgenic Sciences, Inc. (*Reader Service No. 110*). This stabilized tetramethylbenzidine substrate provides, says TSI, high sensitivity and low background for both soluble and precipitating-based diagnostic systems. TMBlue offers the



TSI's TMBlue: a general-purpose substrate for horseradish peroxidase-linked assays.

convenience of a one-bottle system and, as a precipitating reagent, does not fade from immunoblots, RNA/DNA filter hybridizations or antibody-coated beads. The 'precipitating' reagent costs \$75 (US) for 100 ml and the 'soluble' reagent costs \$50 (US) for 200 ml.

Kirkegaard and Perry Laboratories, Inc. has combined affinity purification technology with the sensitivity of the **biotin/streptavidin detection system** to produce six kits for the detection of mouse, rat, rabbit, swine, sheep and goat primary antibodies (*Reader Service No. 111*). The biotin is conjugated to the antibody via a 13-carbon spacer arm, which, the company says, reduces steric hindrance when reacted with enzyme-labelled streptavidin. Non-specific background staining is blocked with normal



Detect immunocomplexes with the Kirkgaard and Perry HistoMark kit.

serum produced in the same animal used to produce the secondary antibody. The immunocomplex can visually be detected using reagents such as HistoMark Orange and Black, diaminobenzidine, 3-Amino-9-ethylcarbazole and 4-Chloro-1-naphthol. The kits comprise normal goat (or rabbit) serum, species-specific link antibody prepared in goat or rabbit, and streptavidin covalently coupled to horseradish peroxidase. Kits are suitable for use with formalin-fixed, paraffin-embedded, or frozen sections, smears and cytospin preparations. Each kit has sufficient reagents for 500 tests.

Antibodies anyone?

Bioproducts for Science, Inc. has introduced a **monoclonal antibody to mature mouse macrophages** (*Reader Service No. 112*). The rat mAb to the mouse F4/80 antigen recognises the 160 kDa plasma-membrane component of mouse mononuclear phagocytes. It can, therefore, be used as a maturation marker for macrophages in bone marrow culture and identified by immunocytochemical methods. Other uses include differentiation, physiology and heterogeneity studies.

BioGenex Laboratories has introduced a mouse monoclonal antibody for the qualitative **localization of muscle-specific actin** in formalin-fixed or paraffin-embedded tissue sections (*Reader Service No. 113*). The mAb HHF35 recognises actin α and γ isotypes of skeletal, cardiac and smooth muscle cells. It is non-reactive with other mesenchymal cells and all epithelial cells except myoepithelium. The mouse mAb can be supplied as a ready-to-use reagent for immunohistochemical staining, or in kit form with Histogen (PAP) or StrAviGen (BSA)

detection systems — sufficient for staining up to 50 slides.

Laboratory Impex Ltd are the UK distributors of T Cell Sciences' **panel of monoclonal antibodies to the variable region epitopes** of the human T-cell antigen receptor (*Reader Service No. 114*). Using the Diversi-T $\alpha\beta$ Screening Panel 1A, which consists of seven antibodies, researchers can investigate several TCR-V regions at the protein level and characterize T-cell subsets based on TCR usage in peripheral blood lymphocytes, cultured thymocytes, T-cell clones, tissue sections, cell lysates and T-cell-derived tumours. The mAbs may have applications in determining the role of T cells in autoimmune diseases, cancer, bacterial and viral infections. They can be used in flow cytometry analysis, immunochemical staining, immunoprecipitation procedures and mitogenic expansion of V-region specific cells. Each mAb is sufficient for 25 flow cytometry tests.

Small-volume filters

Millipore Corp. has introduced Ultrafree-CL **centrifugal filter units** designed to process 2-ml volumes of biological solutions prior to cell-based and instrument analyses (*Reader Service No. 115*). The Ultrafree-CL filter unit comprises a 2-ml sample cup with membrane sealed to the bottom, which then sits inside a 5-ml centrifuge tube. The unit can also be used in standard 15-ml tube fixed-angle centrifugal rotors. Millipore supply the units with low-binding PL series ultrafiltration membranes for the separation of nucleotides from DNA, amino acids from serum proteins, and monosaccharides from starches, or with low-protein-binding microporous Durapore membranes.

The G-Prep centrifuge filter from Gelman Sciences is designed to simultaneously **filter and clarify multiple small-volume samples** using a standard centrifuge (*Reader Service No. 116*). A five-to ten-minute centrifugation is usually all that is required to clarify biologicals, HPLC solvents, water samples and other liquid suspensions with, Gelman says, very low hold-up volumes. G-Prep filtration units consist of a polypropylene centrifuge tube available with either nylon or PTFE membrane filters in 0.2 or 0.45 μ m pore sizes.

Miscellaneous items

Test your knowledge of the celestial planisphere with the 1000-piece **jigsaw** from the Great American Puzzle Factory (*Reader Service No. 117*). According to the company, the jigsaw is a scientifically accurate representation of the northern and southern nighttime skies and measures 20 $\times$ 27 inches when assembled. A special feature of the puzzle is the phos-



Puzzle it out with this educational jigsaw.

phorescent ink that illuminates 88 constellations and the Milky Way galaxy. The jigsaw, costing about \$15 (US), is supplied with a 30-page guidebook to the stars, galaxies, black holes and supernovae.

"Electron Microscopy" is the latest **brochure** from Oxford Instruments for new and existing users of electron microscopes and accessories (*Reader Service No. 118*). The brochure focuses on three main



Oxford Instruments' brochure for EM users.

areas of EM: SEM, TEM and sample preparation for EM, and should be of interest to researchers in a variety of scientific fields.

A new **technical information bulletin** outlining Beckman's EasyLIMS software is now available (*Reader Service No. 119*). The bulletin gives an in-depth overview of EasyLIMS, which is a new LIMS package for IBM PCs that features Microsoft windows. Sections in the bulletin detail the features and benefits of EasyLIMS, the flexibility and user-definability of the package, and the report formats that can be generated.

Dionex Corp. has issued a new **brochure** describing their new line of columns for proteins and nucleic acids (*Reader Service No. 120*). ProPac PA1 columns use MicroBead resin technology and, says Dionex, allow ion exchange separations to be performed with resolutions approaching that of reversed phase. □

These notes are compiled by Diane Gershon from information provided by the manufacturer. To obtain additional information about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

Responding to demographic change

Richard Pearson

Despite the fall in the number of young people in recent years, there is cause to believe that student numbers will be 'holding up' and that employers will develop new approaches to recruitment.

By now there are few people in the United Kingdom who have not heard of the demographic downturn whereby the number of young people falls by a third in the decade mid-1980s to the mid-1990s. Yet 18 months ago less than one in six employers was aware of the scale of the downturn and there were many in higher education who feared that the downturn would result in a similar drop in the number of students. We know now, however, that because of the social class effect, student numbers will be holding up during the 1990s (*Nature* 342, 102; 1989).

But this is not the end of the story. While the trends may be well known, the majority of recruiters are still expecting to continue to recruit the same number of school leavers as before, simply by trying harder in building up school links, marketing and improving pay and prospects; yet they cannot all increase market share. While most areas of higher education are responding by seeking to widen access, only a minority of industrial recruiters are thinking about alternative strategies for dealing with the downturn. Some of the retailers are gearing their recruitment to the over 50s while the age limit for apprentices in British Telecom has been raised to 41 and the entry age for the tax inspectorate has been raised to 55. Other companies are relocating out of the South East of the United Kingdom including some local government departments who are moving specific areas of work such as invoicing and computing away from their home district to easier labour markets.

Local perspective

Many employers do, however, recognize that an essential pre-requisite of any such alternative strategy is a knowledge of the problem and its impact on their organization. In the case of young recruits most school leavers look for and find work in their local labour market and thus the need is for more localized demographic data. New research emphasizes this need for a local perspective by highlighting the extreme variability in the school leaver downturn both between and within regions¹.

In general the largest declines are in the North and West of England; the two major areas with the biggest fall in numbers being Liverpool and Sheffield where the decline exceeds 40 per cent. The lowest are in Suffolk, Berkshire, West Sussex and

Sussex where the decline is under 20 per cent. There are also major differences in the extent to which the numbers recover in the late 1990s; in Liverpool and Sheffield there is little recovery expected in numbers by the end of the century while in the southern counties noted above, the numbers are back to 90 per cent of the 1985 figures and rising further. But even within regions there are dramatic differences, for example, while the decline in Manchester is almost 40 per cent and shows only a small recovery by the year 2000 the average decline across the rest of Lancashire is just under 30 per cent and is almost fully reversed by the year 2000.

Other important differences relate to the academic profiles of the school leavers with, for example, 28 per cent of those in the South East attaining five or more 'O' levels compared with only just over 21 per cent in Yorkshire and Humberside. At a more local level the extremes are even greater from highs of almost 40 per cent in some suburban London boroughs to lows of 15 per cent or less in some parts of the North as well as in some inner London boroughs (where the lowest attainment rate fell below 10 per cent).

These differences reflect many and varied factors of which the social class composition of the parents is one of the most important. As the demographic downturn is concentrated in the lower social classes (III-V), where attainment levels are lowest, the downturn in the number of well qualified leavers is rather less than the average. Most of the downturn is concentrated among those likely to leave school without qualifications; where the fall exceeds 40 per cent. The need for would-be recruiters is therefore to look in detail at particular localities in terms of the changes in both the overall numbers and of differing qualification profiles before embarking on recruitment campaigns.

So much for the school leaver population, what about some of the much discussed alternatives? As we move into the 1990s we find a similar level of ignorance about the changing size of the labour force over the next decade. Half the recruiters last year thought that the fall in the number of young people meant that the size of the whole workforce would fall while one in four recruiters thought it would remain the same. Only one in five were aware that the workforce would grow

by as much as five per cent (equal to or over a million people over the next decade), and that 90 per cent of these additional workers will be women.

Different approaches

Slowly employers are becoming aware of another part of the information jigsaw and the need for more localized data, what then, should they be doing about the changes that affect them? The message for employers in the 90s must be to stop reacting tactically to events and to start to develop a more strategic response. Many recruiters are hanging on in the hope that the problem will go away and are using short-term measures such as more overtime and reducing hiring standards. Others are competing harder for the available pool of talent, but this will become an increasingly difficult and expensive means of maintaining recruitment.

A better approach is to recruit and, if necessary, retrain those from non-traditional backgrounds, be they women or older workers, both of whom will be increasingly available. Others are seeking to reduce wastage which will not only lessen costs by using existing talent, but also reduce the need for replacement recruits. Relocation is an option for many and as indicated above it does not have to mean movement of the whole organization, rather it can focus on critical services and activities. Such moves are increasingly being aided by the reducing cost and effectiveness of information technology and communications networks (*Nature* 338, 98; 1989). The most far-sighted responses will, however, be strategic and be those which seek to reduce the existing demand for recruits by improving the utilization of existing staff by training and redeployment, ensuring that the most appropriate people do the job, and that scarce skills are not wasted, investing in technology and working more effectively with existing staff. As the problems are long-term, the current solutions have to move from the short-term tactical responses to the more strategic if employers are to prosper in the 1990s. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Falmer, Brighton, Sussex BN1 9RF, UK.

1. Waite, R. & Pike, G. *School Leaver Decline and Effective Local Solutions* (IMS, 1989).